Our Ref.: **CL000891**

U.S. PATENT APPLICATION

Inventor(s): Stefan KODET
Gennady V. MERKULOV
Karen A. KETCHUM
Wei SHAO
Chunhua YAN
Valentina DI FRANCESCO
Ellen M. BEASLEY

Invention: ISOLATED HUMAN TRANSPORTER PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS, AND USES THEREOF

CELERA GENOMICS CORPORATION.
45 WEST GUDE DR., C2-4#20
ROCKVILLE, MD 20850
(240) 453-3067
Fax (240)-453-3084

SPECIFICATION

ISOLATED HUMAN TRANSPORTER PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS, AND USES THEREOF

RELATED APPLICATIONS

The present application claims priority to provisional application U.S. Serial No. 60/240,836, filed October 17, 2000 (Atty. Docket CL000891-PROV).

FIELD OF THE INVENTION

The present invention is in the field of transporter proteins that are related to the sodium/calcium exchanger subfamily, recombinant DNA molecules, and protein production. The present invention specifically provides novel peptides and proteins that effect ligand transport and nucleic acid molecules encoding such peptide and protein molecules, all of which are useful in the development of human therapeutics and diagnostic compositions and methods.

BACKGROUND OF THE INVENTION

Transporters

Transporter proteins regulate many different functions of a cell, including cell proliferation, differentiation, and signaling processes, by regulating the flow of molecules such as ions and macromolecules, into and out of cells. Transporters are found in the plasma membranes of virtually every cell in eukaryotic organisms. Transporters mediate a variety of cellular functions including regulation of membrane potentials and absorption and secretion of molecules and ion across cell membranes. When present in intracellular membranes of the Golgi apparatus and endocytic vesicles, transporters, such as chloride channels, also regulate organelle pH. For a review, see Greger, R. (1988) Annu. Rev. Physiol. 50:111-122.

Transporters are generally classified by structure and the type of mode of action. In addition, transporters are sometimes classified by the molecule type that is transported,

for example, sugar transporters, chlorine channels, potassium channels, etc. There may be many classes of channels for transporting a single type of molecule (a detailed review of channel types can be found at Alexander, S.P.H. and J.A. Peters: Receptor and transporter nomenclature supplement. Trends Pharmacol. Sci., Elsevier, pp. 65-68 (1997) and http://www-biology.ucsd.edu/~msaier/transport/titlepage2.html.

The following general classification scheme is known in the art and is followed in the present discoveries.

Channel-type transporters. Transmembrane channel proteins of this class are ubiquitously found in the membranes of all types of organisms from bacteria to higher eukaryotes. Transport systems of this type catalyze facilitated diffusion (by an energy-independent process) by passage through a transmembrane aqueous pore or channel without evidence for a carrier-mediated mechanism. These channel proteins usually consist largely of a-helical spanners, although b-strands may also be present and may even comprise the channel. However, outer membrane porin-type channel proteins are excluded from this class and are instead included in class 9.

Carrier-type transporters. Transport systems are included in this class if they utilize a carrier-mediated process to catalyze uniport (a single species is transported by facilitated diffusion), antiport (two or more species are transported in opposite directions in a tightly coupled process, not coupled to a direct form of energy other than chemiosmotic energy) and/or symport (two or more species are transported together in the same direction in a tightly coupled process, not coupled to a direct form of energy other than chemiosmotic energy).

Pyrophosphate bond hydrolysis-driven active transporters. Transport systems are included in this class if they hydrolyze pyrophosphate or the terminal pyrophosphate bond in ATP or another nucleoside triphosphate to drive the active uptake and/or extrusion of a solute or solutes. The transport protein may or may not be transiently phosphorylated, but the substrate is not phosphorylated.

PEP-dependent, phosphoryl transfer-driven group translocators. Transport systems of the bacterial phosphoenolpyruvate:sugar phosphotransferase system are included in this class. The product of the reaction, derived from extracellular sugar, is a cytoplasmic sugar-phosphate.

Decarboxylation-driven active transporters. Transport systems that drive solute (e.g., ion) uptake or extrusion by decarboxylation of a cytoplasmic substrate are included in this class.

Oxidoreduction-driven active transporters. Transport systems that drive transport of a solute (e.g., an ion) energized by the flow of electrons from a reduced substrate to an oxidized substrate are included in this class.

Light-driven active transporters. Transport systems that utilize light energy to drive transport of a solute (e.g., an ion) are included in this class.

Mechanically-driven active transporters. Transport systems are included in this class if they drive movement of a cell or organelle by allowing the flow of ions (or other solutes) through the membrane down their electrochemical gradients.

Outer-membrane porins (of b-structure). These proteins form transmembrane pores or channels that usually allow the energy independent passage of solutes across a membrane. The transmembrane portions of these proteins consist exclusively of b-strands that form a b-barrel. These porin-type proteins are found in the outer membranes of Gram-negative bacteria, mitochondria and eukaryotic plastids.

Methyltransferase-driven active transporters. A single characterized protein currently falls into this category, the Na+-transporting methyltetrahydromethanopterin:coenzyme M methyltransferase.

Non-ribosome-synthesized channel-forming peptides or peptide-like molecules. These molecules, usually chains of L- and D-amino acids as well as other small molecular building blocks such as lactate, form oligomeric transmembrane ion channels. Voltage may induce channel formation by promoting assembly of the transmembrane channel. These peptides are often made by bacteria and fungi as agents of biological warfare.

Non-Proteinaceous Transport Complexes. Ion conducting substances in biological membranes that do not consist of or are not derived from proteins or peptides fall into this category.

Functionally characterized transporters for which sequence data are lacking. Transporters of particular physiological significance will be included in this category even though a family assignment cannot be made.

Putative transporters in which no family member is an established transporter. Putative transport protein families are grouped under this number and will either be classified elsewhere when the transport function of a member becomes established, or will be eliminated from the TC classification system if the proposed transport function is disproven. These families include a member or members for which a transport function has been suggested, but evidence for such a function is not yet compelling.

Auxiliary transport proteins. Proteins that in some way facilitate transport across one or more biological membranes but do not themselves participate directly in transport are included in this class. These proteins always function in conjunction with one or more transport proteins. They may provide a function connected with energy coupling to transport, play a structural role in complex formation or serve a regulatory function.

Transporters of unknown classification. Transport protein families of unknown classification are grouped under this number and will be classified elsewhere when the transport process and energy coupling mechanism are characterized. These families include at least one member for which a transport function has been established, but either the mode of transport or the energy coupling mechanism is not known.

Ion channels

An important type of transporter is the ion channel. Ion channels regulate many different cell proliferation, differentiation, and signaling processes by regulating the flow of ions into and out of cells. Ion channels are found in the plasma membranes of virtually every cell in eukaryotic organisms. Ion channels mediate a variety of cellular functions including regulation of membrane potentials and absorption and secretion of ion across epithelial membranes. When present in intracellular membranes of the Golgi apparatus and endocytic vesicles, ion channels, such as chloride channels, also regulate organelle pH. For a review, see Greger, R. (1988) Annu. Rev. Physiol. 50:111-122.

Ion channels are generally classified by structure and the type of mode of action. For example, extracellular ligand gated channels (ELGs) are comprised of five polypeptide subunits, with each subunit having 4 membrane spanning domains, and are activated by the binding of an extracellular ligand to the channel. In addition, channels are sometimes classified by the ion type that is transported, for example, chlorine

channels, potassium channels, etc. There may be many classes of channels for transporting a single type of ion (a detailed review of channel types can be found at Alexander, S.P.H. and J.A. Peters (1997). Receptor and ion channel nomenclature supplement. Trends Pharmacol. Sci., Elsevier, pp. 65-68 and http://www-biology.ucsd.edu/~msaier/transport/toc.html.

There are many types of ion channels based on structure. For example, many ion channels fall within one of the following groups: extracellular ligand-gated channels (ELG), intracellular ligand-gated channels (ILG), inward rectifying channels (INR), intercellular (gap junction) channels, and voltage gated channels (VIC). There are additionally recognized other channel families based on ion-type transported, cellular location and drug sensitivity. Detailed information on each of these, their activity, ligand type, ion type, disease association, drugability, and other information pertinent to the present invention, is well known in the art.

Extracellular ligand-gated channels, ELGs, are generally comprised of five polypeptide subunits, Unwin, N. (1993), Cell 72: 31-41; Unwin, N. (1995), Nature 373: 37-43; Hucho, F., et al., (1996) J. Neurochem. 66: 1781-1792; Hucho, F., et al., (1996) Eur. J. Biochem. 239: 539-557; Alexander, S.P.H. and J.A. Peters (1997), Trends Pharmacol. Sci., Elsevier, pp. 4-6; 36-40; 42-44; and Xue, H. (1998) J. Mol. Evol. 47: 323-333. Each subunit has 4 membrane spanning regions: this serves as a means of identifying other members of the ELG family of proteins. ELG bind a ligand and in response modulate the flow of ions. Examples of ELG include most members of the neurotransmitter-receptor family of proteins, e.g., GABAI receptors. Other members of this family of ion channels include glycine receptors, ryandyne receptors, and ligand gated calcium channels.

Sodium/Calcium Exchangers

The protein provided by the present invention is a novel sodium/calcium exchanger. Sodium/calcium exchangers (NCX) rapidly import calcium during excitation impulse. Intracellular calcium concentrations vary greatly during the excitation/relaxation cycle. In contrast, extracellular calcium concentrations are maintained at relatively steady levels, despite wide variations in the amounts of calcium supplied with food.

There are at least three known mammalian NCX genes and a number of alternatively spliced isoforms. NCX sequences are highly conserved. NCX proteins contain 9 transmembrane domains and are regulated by calcium and sodium ions and, to some extent, by phosphorylation.

NCX proteins initiate cardiac myocyte contractions; this effect has been confirmed by *in vitro* experiments. Together with calsequestrin, a calcium binding protein, NCX proteins maintain calcium homeostasis in the heart muscle. This regulatory mechanism depends on the gene dosage, as evident from experiments with transgenic animals. Variations in expression levels of these proteins may be associated with some forms of heart disease.

Calcium transporters can mediate divalent ion toxicity. Barium and strontium can be carried by these channels into the cell, albeit at slower rates than calcium, which is the natural substrate. A panel of bivalent cations, such as copper, lead, cadmium, cobalt and nickel, inhibit calcium flow, but do not penetrate the cell membrane. Bivalent and trivalent iron, manganese, and zinc show no effect.

The sequence of the sodium/calcium exchanger provided by the present invention may be used to screen human populations for mutations associated with neurological conditions and heart disease. Furthermore, drugs can be designed that target this and other transporters.

For a further review of sodium/calcium exchangers, see: Linck *et al.*, *J Pharmacol Exp Ther* 2000 Aug;294(2):648-57; Shen *et al.*, *J Pharmacol Exp Ther* 2000 Aug;294(2):562-70; Philipson *et al.*, *Annu Rev Physiol* 2000;62:111-33; Zhang *et al.*, *Br J Pharmacol* 2000 Jun;130(3):485-8; and Vercesi *et al.*, *FEBS Lett* 2000 May 12;473(2):203-6.

The Voltage-gated Ion Channel (VIC) Superfamily

Proteins of the VIC family are ion-selective channel proteins found in a wide range of bacteria, archaea and eukaryotes Hille, B. (1992), Chapter 9: Structure of channel proteins; Chapter 20: Evolution and diversity. In: Ionic Channels of Excitable Membranes, 2nd Ed., Sinaur Assoc. Inc., Pubs., Sunderland, Massachusetts; Sigworth, F.J. (1993), Quart. Rev. Biophys. 27: 1-40; Salkoff, L. and T. Jegla (1995), Neuron 15:

489-492; Alexander, S.P.H. et al., (1997), Trends Pharmacol. Sci., Elsevier, pp. 76-84; Jan, L.Y. et al., (1997), Annu. Rev. Neurosci. 20: 91-123; Doyle, D.A, et al., (1998) Science 280: 69-77; Terlau, H. and W. Stühmer (1998), Naturwissenschaften 85: 437-444. They are often homo- or heterooligomeric structures with several dissimilar subunits (e.g., a1-a2-d-b Ca^{2+} channels, ab_1b_2 Na^+ channels or $(a)_4$-b K^+ channels), but the channel and the primary receptor is usually associated with the a (or a1) subunit. Functionally characterized members are specific for K^+, Na^+ or Ca^{2+}. The K^+ channels usually consist of homotetrameric structures with each a-subunit possessing six transmembrane spanners (TMSs). The a1 and a subunits of the Ca^{2+} and Na^+ channels, respectively, are about four times as large and possess 4 units, each with 6 TMSs separated by a hydrophilic loop, for a total of 24 TMSs. These large channel proteins form heterotetra-unit structures equivalent to the homotetrameric structures of most K^+ channels. All four units of the Ca^{2+} and Na^+ channels are homologous to the single unit in the homotetrameric K^+ channels. Ion flux via the eukaryotic channels is generally controlled by the transmembrane electrical potential (hence the designation, voltage-sensitive) although some are controlled by ligand or receptor binding.

Several putative K^+-selective channel proteins of the VIC family have been identified in prokaryotes. The structure of one of them, the KcsA K^+ channel of *Streptomyces lividans*, has been solved to 3.2 Å resolution. The protein possesses four identical subunits, each with two transmembrane helices, arranged in the shape of an inverted teepee or cone. The cone cradles the "selectivity filter" P domain in its outer end. The narrow selectivity filter is only 12 Å long, whereas the remainder of the channel is wider and lined with hydrophobic residues. A large water-filled cavity and helix dipoles stabilize K^+ in the pore. The selectivity filter has two bound K^+ ions about 7.5 Å apart from each other. Ion conduction is proposed to result from a balance of electrostatic attractive and repulsive forces.

In eukaryotes, each VIC family channel type has several subtypes based on pharmacological and electrophysiological data. Thus, there are five types of Ca^{2+} channels (L, N, P, Q and T). There are at least ten types of K^+ channels, each responding in different ways to different stimuli: voltage-sensitive [Ka, Kv, Kvr, Kvs and Ksr], Ca^{2+}-sensitive [BK_{Ca}, IK_{Ca} and SK_{Ca}] and receptor-coupled [K_M and K_{ACh}]. There are at least

six types of Na^+ channels (I, II, III, μ1, H1 and PN3). Tetrameric channels from both prokaryotic and eukaryotic organisms are known in which each a-subunit possesses 2 TMSs rather than 6, and these two TMSs are homologous to TMSs 5 and 6 of the six TMS unit found in the voltage-sensitive channel proteins. KcsA of *S. lividans* is an example of such a 2 TMS channel protein. These channels may include the K_{Na} (Na^+-activated) and K_{Vol} (cell volume-sensitive) K^+ channels, as well as distantly related channels such as the Tok1 K^+ channel of yeast, the TWIK-1 inward rectifier K^+ channel of the mouse and the TREK-1 K^+ channel of the mouse. Because of insufficient sequence similarity with proteins of the VIC family, inward rectifier K^+ IRK channels (ATP-regulated; G-protein-activated) which possess a P domain and two flanking TMSs are placed in a distinct family. However, substantial sequence similarity in the P region suggests that they are homologous. The b, g and d subunits of VIC family members, when present, frequently play regulatory roles in channel activation/deactivation.

The Epithelial Na^+ Channel (ENaC) Family

The ENaC family consists of over twenty-four sequenced proteins (Canessa, C.M., et al., (1994), Nature 367: 463-467, Le, T. and M.H. Saier, Jr. (1996), Mol. Membr. Biol. 13: 149-157; Garty, H. and L.G. Palmer (1997), Physiol. Rev. 77: 359-396; Waldmann, R., et al., (1997), Nature 386: 173-177; Darboux, I., et al., (1998), J. Biol. Chem. 273: 9424-9429; Firsov, D., et al., (1998), EMBO J. 17: 344-352; Horisberger, J.-D. (1998). Curr. Opin. Struc. Biol. 10: 443-449). All are from animals with no recognizable homologues in other eukaryotes or bacteria. The vertebrate ENaC proteins from epithelial cells cluster tightly together on the phylogenetic tree: voltage-insensitive ENaC homologues are also found in the brain. Eleven sequenced *C. elegans* proteins, including the degenerins, are distantly related to the vertebrate proteins as well as to each other. At least some of these proteins form part of a mechano-transducing complex for touch sensitivity. The homologous *Helix aspersa* (FMRF-amide)-activated Na^+ channel is the first peptide neurotransmitter-gated ionotropic receptor to be sequenced.

Protein members of this family all exhibit the same apparent topology, each with N- and C-termini on the inside of the cell, two amphipathic transmembrane spanning segments, and a large extracellular loop. The extracellular domains contain numerous highly conserved cysteine residues. They are proposed to serve a receptor function.

Mammalian ENaC is important for the maintenance of Na^+ balance and the regulation of blood pressure. Three homologous ENaC subunits, alpha, beta, and gamma, have been shown to assemble to form the highly Na^+-selective channel. The stoichiometry of the three subunits is $alpha_2$, beta1, gamma1 in a heterotetrameric architecture.

The Glutamate-gated Ion Channel (GIC) Family of Neurotransmitter Receptors

Members of the GIC family are heteropentameric complexes in which each of the 5 subunits is of 800-1000 amino acyl residues in length (Nakanishi, N., et al, (1990), Neuron 5: 569-581; Unwin, N. (1993), Cell 72: 31-41; Alexander, S.P.H. and J.A. Peters (1997) Trends Pharmacol. Sci., Elsevier, pp. 36-40). These subunits may span the membrane three or five times as putative a-helices with the N-termini (the glutamate-binding domains) localized extracellularly and the C-termini localized cytoplasmically. They may be distantly related to the ligand-gated ion channels, and if so, they may possess substantial b-structure in their transmembrane regions. However, homology between these two families cannot be established on the basis of sequence comparisons alone. The subunits fall into six subfamilies: a, b, g, d, e and z.

The GIC channels are divided into three types: (1) a-amino-3-hydroxy-5-methyl-4-isoxazole propionate (AMPA)-, (2) kainate- and (3) N-methyl-D-aspartate (NMDA)-selective glutamate receptors. Subunits of the AMPA and kainate classes exhibit 35-40% identity with each other while subunits of the NMDA receptors exhibit 22-24% identity with the former subunits. They possess large N-terminal, extracellular glutamate-binding domains that are homologous to the periplasmic glutamine and glutamate receptors of ABC-type uptake permeases of Gram-negative bacteria. All known members of the GIC family are from animals. The different channel (receptor) types exhibit distinct ion selectivities and conductance properties. The NMDA-selective large conductance channels are highly permeable to monovalent cations and Ca^{2+}. The AMPA- and kainate-selective ion channels are permeable primarily to monovalent cations with only low permeability to Ca^{2+}.

The Chloride Channel (ClC) Family

The ClC family is a large family consisting of dozens of sequenced proteins derived from Gram-negative and Gram-positive bacteria, cyanobacteria, archaea, yeast, plants and animals (Steinmeyer, K., et al., (1991), Nature 354: 301-304; Uchida, S., et al., (1993), J. Biol. Chem. 268: 3821-3824; Huang, M.-E., et al., (1994), J. Mol. Biol. 242: 595-598; Kawasaki, M., et al, (1994), Neuron 12: 597-604; Fisher, W.E., et al., (1995), Genomics. 29:598-606; and Foskett, J.K. (1998), Annu. Rev. Physiol. 60: 689-717). These proteins are essentially ubiquitous, although they are not encoded within genomes of *Haemophilus influenzae*, *Mycoplasma genitalium,* and *Mycoplasma pneumoniae*. Sequenced proteins vary in size from 395 amino acyl residues *(M. jannaschii)* to 988 residues (man). Several organisms contain multiple ClC family paralogues. For example, *Synechocystis* has two paralogues, one of 451 residues in length and the other of 899 residues. *Arabidopsis thaliana* has at least four sequenced paralogues, (775-792 residues), humans also have at least five paralogues (820-988 residues), and *C. elegans* also has at least five (810-950 residues). There are nine known members in mammals, and mutations in three of the corresponding genes cause human diseases. *E. coli*, *Methanococcus jannaschii* and *Saccharomyces cerevisiae* only have one ClC family member each. With the exception of the larger *Synechocystis* paralogue, all bacterial proteins are small (395-492 residues) while all eukaryotic proteins are larger (687-988 residues). These proteins exhibit 10-12 putative transmembrane a-helical spanners (TMSs) and appear to be present in the membrane as homodimers. While one member of the family, *Torpedo* ClC-O, has been reported to have two channels, one per subunit, others are believed to have just one.

All functionally characterized members of the ClC family transport chloride, some in a voltage-regulated process. These channels serve a variety of physiological functions (cell volume regulation; membrane potential stabilization; signal transduction; transepithelial transport, etc.). Different homologues in humans exhibit differing anion selectivities, i.e., ClC4 and ClC5 share a $NO_3^- > Cl^- > Br^- > I^-$ conductance sequence, while ClC3 has an $I^- > Cl^-$ selectivity. The ClC4 and ClC5 channels and others exhibit outward rectifying currents with currents only at voltages more positive than +20mV.

Animal Inward Rectifier K^+ Channel (IRK-C) Family

IRK channels possess the "minimal channel-forming structure" with only a P domain, characteristic of the channel proteins of the VIC family, and two flanking transmembrane spanners (Shuck, M.E., et al., (1994), J. Biol. Chem. 269: 24261-24270; Ashen, M.D., et al., (1995), Am. J. Physiol. 268: H506-H511; Salkoff, L. and T. Jegla (1995), Neuron 15: 489-492; Aguilar-Bryan, L., et al., (1998), Physiol. Rev. 78: 227-245; Ruknudin, A., et al., (1998), J. Biol. Chem. 273: 14165-14171). They may exist in the membrane as homo- or heterooligomers. They have a greater tendency to let K^+ flow into the cell than out. Voltage-dependence may be regulated by external K^+, by internal Mg^{2+}, by internal ATP and/or by G-proteins. The P domains of IRK channels exhibit limited sequence similarity to those of the VIC family, but this sequence similarity is insufficient to establish homology. Inward rectifiers play a role in setting cellular membrane potentials, and the closing of these channels upon depolarization permits the occurrence of long duration action potentials with a plateau phase. Inward rectifiers lack the intrinsic voltage sensing helices found in VIC family channels. In a few cases, those of Kir1.1a and Kir6.2, for example, direct interaction with a member of the ABC superfamily has been proposed to confer unique functional and regulatory properties to the heteromeric complex, including sensitivity to ATP. The SUR1 sulfonylurea receptor (spQ09428) is the ABC protein that regulates the Kir6.2 channel in response to ATP, and CFTR may regulate Kir1.1a. Mutations in SUR1 are the cause of familial persistent hyperinsulinemic hypoglycemia in infancy (PHHI), an autosomal recessive disorder characterized by unregulated insulin secretion in the pancreas.

ATP-gated Cation Channel (ACC) Family

Members of the ACC family (also called P2X receptors) respond to ATP, a functional neurotransmitter released by exocytosis from many types of neurons (North, R.A. (1996), Curr. Opin. Cell Biol. 8: 474-483; Soto, F., M. Garcia-Guzman and W. Stühmer (1997), J. Membr. Biol. 160: 91-100). They have been placed into seven groups ($P2X_1$ - $P2X_7$) based on their pharmacological properties. These channels, which function at neuron-neuron and neuron-smooth muscle junctions, may play roles in the control of

blood pressure and pain sensation. They may also function in lymphocyte and platelet physiology. They are found only in animals.

The proteins of the ACC family are quite similar in sequence (>35% identity), but they possess 380-1000 amino acyl residues per subunit with variability in length localized primarily to the C-terminal domains. They possess two transmembrane spanners, one about 30-50 residues from their N-termini, the other near residues 320-340. The extracellular receptor domains between these two spanners (of about 270 residues) are well conserved with numerous conserved glycyl and cysteyl residues. The hydrophilic C-termini vary in length from 25 to 240 residues. They resemble the topologically similar epithelial Na^+ channel (ENaC) proteins in possessing (a) N- and C-termini localized intracellularly, (b) two putative transmembrane spanners, (c) a large extracellular loop domain, and (d) many conserved extracellular cysteyl residues. ACC family members are, however, not demonstrably homologous with them. ACC channels are probably hetero- or homomultimers and transport small monovalent cations (Me^+). Some also transport Ca^{2+}; a few also transport small metabolites.

The Ryanodine-Inositol 1,4,5-triphosphate Receptor Ca^{2+} Channel (RIR-CaC) Family

Ryanodine (Ry)-sensitive and inositol 1,4,5-triphosphate (IP3)-sensitive Ca^{2+}-release channels function in the release of Ca^{2+} from intracellular storage sites in animal cells and thereby regulate various Ca^{2+} -dependent physiological processes (Hasan, G. et al., (1992) Development 116: 967-975; Michikawa, T., et al., (1994), J. Biol. Chem. 269: 9184-9189; Tunwell, R.E.A., (1996), Biochem. J. 318: 477-487; Lee, A.G. (1996) *Biomembranes,* Vol. 6, Transmembrane Receptors and Channels (A.G. Lee, ed.), JAI Press, Denver, CO., pp 291-326; Mikoshiba, K., et al., (1996) J. Biochem. Biomem. 6: 273-289). Ry receptors occur primarily in muscle cell sarcoplasmic reticular (SR) membranes, and IP3 receptors occur primarily in brain cell endoplasmic reticular (ER) membranes where they effect release of Ca^{2+} into the cytoplasm upon activation (opening) of the channel.

The Ry receptors are activated as a result of the activity of dihydropyridine-sensitive Ca^{2+} channels. The latter are members of the voltage-sensitive ion channel

(VIC) family. Dihydropyridine-sensitive channels are present in the T-tubular systems of muscle tissues.

Ry receptors are homotetrameric complexes with each subunit exhibiting a molecular size of over 500,000 daltons (about 5,000 amino acyl residues). They possess C-terminal domains with six putative transmembrane a -helical spanners (TMSs). Putative pore-forming sequences occur between the fifth and sixth TMSs as suggested for members of the VIC family. The large N-terminal hydrophilic domains and the small C-terminal hydrophilic domains are localized to the cytoplasm. Low resolution 3-dimensional structural data are available. Mammals possess at least three isoforms that probably arose by gene duplication and divergence before divergence of the mammalian species. Homologues are present in humans and *Caenorabditis elegans.*

IP_3 receptors resemble Ry receptors in many respects. (1) They are homotetrameric complexes with each subunit exhibiting a molecular size of over 300,000 daltons (about 2,700 amino acyl residues). (2) They possess C-terminal channel domains that are homologous to those of the Ry receptors. (3) The channel domains possess six putative TMSs and a putative channel lining region between TMSs 5 and 6. (4) Both the large N-terminal domains and the smaller C-terminal tails face the cytoplasm. (5) They possess covalently linked carbohydrate on extracytoplasmic loops of the channel domains. (6) They have three currently recognized isoforms (types 1, 2, and 3) in mammals which are subject to differential regulation and have different tissue distributions.

IP_3 receptors possess three domains: N-terminal IP_3-binding domains, central coupling or regulatory domains and C-terminal channel domains. Channels are activated by IP_3 binding, and like the Ry receptors, the activities of the IP_3 receptor channels are regulated by phosphorylation of the regulatory domains, catalyzed by various protein kinases. They predominate in the endoplasmic reticular membranes of various cell types in the brain but have also been found in the plasma membranes of some nerve cells derived from a variety of tissues.

The channel domains of the Ry and IP_3 receptors comprise a coherent family that in spite of apparent structural similarities, do not show appreciable sequence similarity of the proteins of the VIC family. The Ry receptors and the IP_3 receptors cluster separately

on the RIR-CaC family tree. They both have homologues in *Drosophila.* Based on the phylogenetic tree for the family, the family probably evolved in the following sequence: (1) A gene duplication event occurred that gave rise to Ry and IP_3 receptors in invertebrates. (2) Vertebrates evolved from invertebrates. (3) The three isoforms of each receptor arose as a result of two distinct gene duplication events. (4) These isoforms were transmitted to mammals before divergence of the mammalian species.

The Organellar Chloride Channel (O-ClC) Family

Proteins of the O-ClC family are voltage-sensitive chloride channels found in intracellular membranes but not the plasma membranes of animal cells (Landry, D, et al., (1993), J. Biol. Chem. 268: 14948-14955; Valenzuela, Set al., (1997), J. Biol. Chem. 272: 12575-12582; and Duncan, R.R., et al., (1997), J. Biol. Chem. 272: 23880-23886).

They are found in human nuclear membranes, and the bovine protein targets to the microsomes, but not the plasma membrane, when expressed in *Xenopus laevis* oocytes. These proteins are thought to function in the regulation of the membrane potential and in transepithelial ion absorption and secretion in the kidney. They possess two putative transmembrane a-helical spanners (TMSs) with cytoplasmic N- and C-termini and a large luminal loop that may be glycosylated. The bovine protein is 437 amino acyl residues in length and has the two putative TMSs at positions 223-239 and 367-385. The human nuclear protein is much smaller (241 residues). A *C. elegans* homologue is 260 residues long.

Transporter proteins, particularly members of the sodium/calcium exchanger subfamily, are a major target for drug action and development. Accordingly, it is valuable to the field of pharmaceutical development to identify and characterize previously unknown transport proteins. The present invention advances the state of the art by providing previously unidentified human transport proteins.

SUMMARY OF THE INVENTION

The present invention is based in part on the identification of amino acid sequences of human transporter peptides and proteins that are related to the sodium/calcium exchanger subfamily, as well as allelic variants and other mammalian

orthologs thereof. These unique peptide sequences, and nucleic acid sequences that encode these peptides, can be used as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins, and serve as targets for the development of human therapeutic agents that modulate transporter activity in cells and tissues that express the transporter. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

DESCRIPTION OF THE FIGURE SHEETS

FIGURE 1 provides the nucleotide sequence of a cDNA molecule or transcript sequence that encodes the transporter protein of the present invention (SEQ ID NO:1). In addition structure and functional information is provided, such as ATG start, stop and tissue distribution, where available, that allows one to readily determine specific uses of inventions based on this molecular sequence. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

FIGURE 2 provides the predicted amino acid sequence of the transporter of the present invention. (SEQ ID NO:2) In addition structure and functional information such as protein family, function, and modification sites is provided where available, allowing one to readily determine specific uses of inventions based on this molecular sequence.

FIGURE 3 provides genomic sequences that span the gene encoding the transporter protein of the present invention (SEQ ID NO: 3). In addition structure and functional information, such as intron/exon structure, promoter location, etc., is provided where available, allowing one to readily determine specific uses of inventions based on this molecular sequence. 140 SNPs, including 6 indels, have been identified in the gene encoding the transporter protein provided by the present invention and are given in Figure 3.

DETAILED DESCRIPTION OF THE INVENTION

General Description

The present invention is based on the sequencing of the human genome. During the sequencing and assembly of the human genome, analysis of the sequence information revealed previously unidentified fragments of the human genome that encode peptides that share structural and/or sequence homology to protein/peptide/domains identified and characterized within the art as being a transporter protein or part of a transporter protein and are related to the sodium/calcium exchanger subfamily. Utilizing these sequences, additional genomic sequences were assembled and transcript and/or cDNA sequences were isolated and characterized. Based on this analysis, the present invention provides amino acid sequences of human transporter peptides and proteins that are related to the sodium/calcium exchanger subfamily, nucleic acid sequences in the form of transcript sequences, cDNA sequences and/or genomic sequences that encode these transporter peptides and proteins, nucleic acid variation (allelic information), tissue distribution of expression, and information about the closest art known protein/peptide/domain that has structural or sequence homology to the transporter of the present invention.

In addition to being previously unknown, the peptides that are provided in the present invention are selected based on their ability to be used for the development of commercially important products and services. Specifically, the present peptides are selected based on homology and/or structural relatedness to known transporter proteins of the sodium/calcium exchanger subfamily and the expression pattern observed . Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.. The art has clearly established the commercial importance of members of this family of proteins and proteins that have expression patterns similar to that of the present gene. Some of the more specific features of the peptides of the present invention, and the uses thereof, are described herein, particularly in the Background of the Invention and in the annotation provided in the Figures, and/or are known within the art for each of the known sodium/calcium exchanger family or subfamily of transporter proteins.

Specific Embodiments

Peptide Molecules

The present invention provides nucleic acid sequences that encode protein molecules that have been identified as being members of the transporter family of proteins and are related to the sodium/calcium exchanger subfamily (protein sequences are provided in Figure 2, transcript/cDNA sequences are provided in Figures 1 and genomic sequences are provided in Figure 3). The peptide sequences provided in Figure 2, as well as the obvious variants described herein, particularly allelic variants as identified herein and using the information in Figure 3, will be referred herein as the transporter peptides of the present invention, transporter peptides, or peptides/proteins of the present invention.

The present invention provides isolated peptide and protein molecules that consist of, consist essentially of, or comprising the amino acid sequences of the transporter peptides disclosed in the Figure 2, (encoded by the nucleic acid molecule shown in Figure 1, transcript/cDNA or Figure 3, genomic sequence), as well as all obvious variants of these peptides that are within the art to make and use. Some of these variants are described in detail below.

As used herein, a peptide is said to be "isolated" or "purified" when it is substantially free of cellular material or free of chemical precursors or other chemicals. The peptides of the present invention can be purified to homogeneity or other degrees of purity. The level of purification will be based on the intended use. The critical feature is that the preparation allows for the desired function of the peptide, even if in the presence of considerable amounts of other components (the features of an isolated nucleic acid molecule is discussed below).

In some uses, "substantially free of cellular material" includes preparations of the peptide having less than about 30% (by dry weight) other proteins (i.e., contaminating protein), less than about 20% other proteins, less than about 10% other proteins, or less than about 5% other proteins. When the peptide is recombinantly produced, it can also be substantially free of culture medium, i.e., culture medium represents less than about 20% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of the peptide in which it is separated from chemical precursors or other chemicals that are involved in its synthesis. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of the transporter peptide having less than about 30% (by dry weight) chemical precursors or other chemicals, less than about 20% chemical precursors or other chemicals, less than about 10% chemical precursors or other chemicals, or less than about 5% chemical precursors or other chemicals.

The isolated transporter peptide can be purified from cells that naturally express it, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. For example, a nucleic acid molecule encoding the transporter peptide is cloned into an expression vector, the expression vector introduced into a host cell and the protein expressed in the host cell. The protein can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques. Many of these techniques are described in detail below.

Accordingly, the present invention provides proteins that consist of the amino acid sequences provided in Figure 2 (SEQ ID NO:2), for example, proteins encoded by the transcript/cDNA nucleic acid sequences shown in Figure 1 (SEQ ID NO:1) and the genomic sequences provided in Figure 3 (SEQ ID NO:3). The amino acid sequence of such a protein is provided in Figure 2. A protein consists of an amino acid sequence when the amino acid sequence is the final amino acid sequence of the protein.

The present invention further provides proteins that consist essentially of the amino acid sequences provided in Figure 2 (SEQ ID NO:2), for example, proteins encoded by the transcript/cDNA nucleic acid sequences shown in Figure 1 (SEQ ID NO:1) and the genomic sequences provided in Figure 3 (SEQ ID NO:3). A protein consists essentially of an amino acid sequence when such an amino acid sequence is present with only a few additional amino acid residues, for example from about 1 to about 100 or so additional residues, typically from 1 to about 20 additional residues in the final protein.

The present invention further provides proteins that comprise the amino acid sequences provided in Figure 2 (SEQ ID NO:2), for example, proteins encoded by the

transcript/cDNA nucleic acid sequences shown in Figure 1 (SEQ ID NO:1) and the genomic sequences provided in Figure 3 (SEQ ID NO:3). A protein comprises an amino acid sequence when the amino acid sequence is at least part of the final amino acid sequence of the protein. In such a fashion, the protein can be only the peptide or have additional amino acid molecules, such as amino acid residues (contiguous encoded sequence) that are naturally associated with it or heterologous amino acid residues/peptide sequences. Such a protein can have a few additional amino acid residues or can comprise several hundred or more additional amino acids. The preferred classes of proteins that are comprised of the transporter peptides of the present invention are the naturally occurring mature proteins. A brief description of how various types of these proteins can be made/isolated is provided below.

The transporter peptides of the present invention can be attached to heterologous sequences to form chimeric or fusion proteins. Such chimeric and fusion proteins comprise a transporter peptide operatively linked to a heterologous protein having an amino acid sequence not substantially homologous to the transporter peptide. "Operatively linked" indicates that the transporter peptide and the heterologous protein are fused in-frame. The heterologous protein can be fused to the N-terminus or C-terminus of the transporter peptide.

In some uses, the fusion protein does not affect the activity of the transporter peptide *per se*. For example, the fusion protein can include, but is not limited to, enzymatic fusion proteins, for example beta-galactosidase fusions, yeast two-hybrid GAL fusions, poly-His fusions, MYC-tagged, HI-tagged and Ig fusions. Such fusion proteins, particularly poly-His fusions, can facilitate the purification of recombinant transporter peptide. In certain host cells (e.g., mammalian host cells), expression and/or secretion of a protein can be increased by using a heterologous signal sequence.

A chimeric or fusion protein can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different protein sequences are ligated together in-frame in accordance with conventional techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two

consecutive gene fragments which can subsequently be annealed and re-amplified to generate a chimeric gene sequence (see Ausubel *et al.*, *Current Protocols in Molecular Biology*, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST protein). A transporter peptide-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the transporter peptide.

As mentioned above, the present invention also provides and enables obvious variants of the amino acid sequence of the proteins of the present invention, such as naturally occurring mature forms of the peptide, allelic/sequence variants of the peptides, non-naturally occurring recombinantly derived variants of the peptides, and orthologs and paralogs of the peptides. Such variants can readily be generated using art-known techniques in the fields of recombinant nucleic acid technology and protein biochemistry. It is understood, however, that variants exclude any amino acid sequences disclosed prior to the invention.

Such variants can readily be identified/made using molecular techniques and the sequence information disclosed herein. Further, such variants can readily be distinguished from other peptides based on sequence and/or structural homology to the transporter peptides of the present invention. The degree of homology/identity present will be based primarily on whether the peptide is a functional variant or non-functional variant, the amount of divergence present in the paralog family and the evolutionary distance between the orthologs.

To determine the percent identity of two amino acid sequences or two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, at least 30%, 40%, 50%, 60%, 70%, 80%, or 90% or more of a reference sequence is aligned for comparison purposes. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or

nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity and similarity between two sequences can be accomplished using a mathematical algorithm. (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part 1*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (Devereux, J., *et al., Nucleic Acids Res. 12(1)*:387 (1984)) (available at http://www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Myers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against sequence databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (*J. Mol. Biol.* 215:403-10 (1990)). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences

homologous to the nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to the proteins of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (*Nucleic Acids Res.* 25(17):3389-3402 (1997)). When utilizing BLAST and gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used.

Full-length pre-processed forms, as well as mature processed forms, of proteins that comprise one of the peptides of the present invention can readily be identified as having complete sequence identity to one of the transporter peptides of the present invention as well as being encoded by the same genetic locus as the transporter peptide provided herein. As indicated by the data presented in Figure 3, the map position was determined to be on chromosome 14 by ePCR.

Allelic variants of a transporter peptide can readily be identified as being a human protein having a high degree (significant) of sequence homology/identity to at least a portion of the transporter peptide as well as being encoded by the same genetic locus as the transporter peptide provided herein. Genetic locus can readily be determined based on the genomic information provided in Figure 3, such as the genomic sequence mapped to the reference human. As indicated by the data presented in Figure 3, the map position was determined to be on chromosome 14 by ePCR. As used herein, two proteins (or a region of the proteins) have significant homology when the amino acid sequences are typically at least about 70-80%, 80-90%, and more typically at least about 90-95% or more homologous. A significantly homologous amino acid sequence, according to the present invention, will be encoded by a nucleic acid sequence that will hybridize to a transporter peptide encoding nucleic acid molecule under stringent conditions as more fully described below.

Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 140 SNP variants were found, including 6 indels (indicated by a "-") and 1 SNPs in exons. The others were found in in introns and regions 5' and 3' of the ORF. Such SNPs in introns and outside the ORF may affect control/regulatory elements.

Paralogs of a transporter peptide can readily be identified as having some degree of significant sequence homology/identity to at least a portion of the transporter peptide, as being encoded by a gene from humans, and as having similar activity or function. Two proteins will typically be considered paralogs when the amino acid sequences are typically at least about 60% or greater, and more typically at least about 70% or greater homology through a given region or domain. Such paralogs will be encoded by a nucleic acid sequence that will hybridize to a transporter peptide encoding nucleic acid molecule under moderate to stringent conditions as more fully described below.

Orthologs of a transporter peptide can readily be identified as having some degree of significant sequence homology/identity to at least a portion of the transporter peptide as well as being encoded by a gene from another organism. Preferred orthologs will be isolated from mammals, preferably primates, for the development of human therapeutic targets and agents. Such orthologs will be encoded by a nucleic acid sequence that will hybridize to a transporter peptide encoding nucleic acid molecule under moderate to stringent conditions, as more fully described below, depending on the degree of relatedness of the two organisms yielding the proteins.

Non-naturally occurring variants of the transporter peptides of the present invention can readily be generated using recombinant techniques. Such variants include, but are not limited to deletions, additions and substitutions in the amino acid sequence of the transporter peptide. For example, one class of substitutions are conserved amino acid substitution. Such substitutions are those that substitute a given amino acid in a transporter peptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu, and Ile; interchange of the hydroxyl residues Ser and Thr; exchange of the acidic residues Asp and Glu; substitution between the amide residues Asn and Gln; exchange of the basic residues Lys and Arg; and replacements among the aromatic residues Phe and Tyr. Guidance concerning which amino acid changes are likely to be phenotypically silent are found in Bowie *et al.*, *Science 247*:1306-1310 (1990).

Variant transporter peptides can be fully functional or can lack function in one or more activities, e.g. ability to bind ligand, ability to transport ligand, ability to mediate signaling, etc. Fully functional variants typically contain only conservative variation or

variation in non-critical residues or in non-critical regions. Figure 2 provides the result of protein analysis and can be used to identify critical domains/regions. Functional variants can also contain substitution of similar amino acids that result in no change or an insignificant change in function. Alternatively, such substitutions may positively or negatively affect function to some degree.

Non-functional variants typically contain one or more non-conservative amino acid substitutions, deletions, insertions, inversions, or truncation or a substitution, insertion, inversion, or deletion in a critical residue or critical region.

Amino acids that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham *et al.*, *Science 244*:1081-1085 (1989)), particularly using the results provided in Figure 2. The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as transporter activity or in assays such as an *in vitro* proliferative activity. Sites that are critical for binding partner/substrate binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith *et al., J. Mol. Biol. 224*:899-904 (1992); de Vos *et al. Science 255*:306-312 (1992)).

The present invention further provides fragments of the transporter peptides, in addition to proteins and peptides that comprise and consist of such fragments, particularly those comprising the residues identified in Figure 2. The fragments to which the invention pertains, however, are not to be construed as encompassing fragments that may be disclosed publicly prior to the present invention.

As used herein, a fragment comprises at least 8, 10, 12, 14, 16, or more contiguous amino acid residues from a transporter peptide. Such fragments can be chosen based on the ability to retain one or more of the biological activities of the transporter peptide or could be chosen for the ability to perform a function, e.g. bind a substrate or act as an immunogen. Particularly important fragments are biologically active fragments, peptides that are, for example, about 8 or more amino acids in length. Such fragments will typically comprise a domain or motif of the transporter peptide, e.g., active site, a transmembrane domain or a substrate-binding domain. Further, possible fragments include, but are not limited to, domain or motif containing fragments, soluble peptide fragments, and fragments containing

immunogenic structures. Predicted domains and functional sites are readily identifiable by computer programs well known and readily available to those of skill in the art (e.g., PROSITE analysis). The results of one such analysis are provided in Figure 2.

Polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids. Further, many amino acids, including the terminal amino acids, may be modified by natural processes, such as processing and other post-translational modifications, or by chemical modification techniques well known in the art. Common modifications that occur naturally in transporter peptides are described in basic texts, detailed monographs, and the research literature, and they are well known to those of skill in the art (some of these features are identified in Figure 2).

Known modifications include, but are not limited to, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks, formation of cystine, formation of pyroglutamate, formylation, gamma carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well known to those of skill in the art and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as *Proteins - Structure and Molecular Properties*, 2nd Ed., T.E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as by Wold, F., *Posttranslational Covalent Modification of Proteins*, B.C. Johnson, Ed., Academic Press, New York 1-12 (1983); Seifter *et al.* (*Meth. Enzymol. 182*: 626-646 (1990)) and Rattan *et al.* (*Ann. N.Y. Acad. Sci. 663*:48-62 (1992)).

Accordingly, the transporter peptides of the present invention also encompass derivatives or analogs in which a substituted amino acid residue is not one encoded by the

genetic code, in which a substituent group is included, in which the mature transporter peptide is fused with another compound, such as a compound to increase the half-life of the transporter peptide (for example, polyethylene glycol), or in which the additional amino acids are fused to the mature transporter peptide, such as a leader or secretory sequence or a sequence for purification of the mature transporter peptide or a pro-protein sequence.

Protein/Peptide Uses

The proteins of the present invention can be used in substantial and specific assays related to the functional information provided in the Figures; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its binding partner or ligand) in biological fluids; and as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state). Where the protein binds or potentially binds to another protein or ligand (such as, for example, in a transporter-effector protein interaction or transporter-ligand interaction), the protein can be used to identify the binding partner/ligand so as to develop a system to identify inhibitors of the binding interaction. Any or all of these uses are capable of being developed into reagent grade or kit format for commercialization as commercial products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

The potential uses of the peptides of the present invention are based primarily on the source of the protein as well as the class/action of the protein. For example, transporters isolated from humans and their human/mammalian orthologs serve as targets for identifying agents for use in mammalian therapeutic applications, e.g. a human drug, particularly in modulating a biological or pathological response in a cell or tissue that expresses the transporter. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in

the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. A large percentage of pharmaceutical agents are being developed that modulate the activity of transporter proteins, particularly members of the sodium/calcium exchanger subfamily (see Background of the Invention). The structural and functional information provided in the Background and Figures provide specific and substantial uses for the molecules of the present invention, particularly in combination with the expression information provided in Figure 1. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. Such uses can readily be determined using the information provided herein, that known in the art and routine experimentation.

The proteins of the present invention (including variants and fragments that may have been disclosed prior to the present invention) are useful for biological assays related to transporters that are related to members of the sodium/calcium exchanger subfamily. Such assays involve any of the known transporter functions or activities or properties useful for diagnosis and treatment of transporter-related conditions that are specific for the subfamily of transporters that the one of the present invention belongs to, particularly in cells and tissues that express the transporter. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

The proteins of the present invention are also useful in drug screening assays, in cell-based or cell-free systems ((Hodgson, Bio/technology, 1992, Sept 10(9);973-80). Cell-based systems can be native, i.e., cells that normally express the transporter, as a biopsy or expanded in cell culture. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. In an alternate embodiment, cell-based assays involve recombinant host cells expressing the transporter protein.

The polypeptides can be used to identify compounds that modulate transporter activity of the protein in its natural state or an altered form that causes a specific disease or pathology associated with the transporter. Both the transporters of the present invention and appropriate variants and fragments can be used in high-throughput screens to assay candidate compounds for the ability to bind to the transporter. These compounds can be further screened against a functional transporter to determine the effect of the compound on the transporter activity. Further, these compounds can be tested in animal or invertebrate systems to determine activity/effectiveness. Compounds can be identified that activate (agonist) or inactivate (antagonist) the transporter to a desired degree.

Further, the proteins of the present invention can be used to screen a compound for the ability to stimulate or inhibit interaction between the transporter protein and a molecule that normally interacts with the transporter protein, e.g. a substrate or a component of the signal pathway that the transporter protein normally interacts (for example, another transporter). Such assays typically include the steps of combining the transporter protein with a candidate compound under conditions that allow the transporter protein, or fragment, to interact with the target molecule, and to detect the formation of a complex between the protein and the target or to detect the biochemical consequence of the interaction with the transporter protein and the target, such as any of the associated effects of signal transduction such as changes in membrane potential, protein phosphorylation, cAMP turnover, and adenylate cyclase activation, etc.

Candidate compounds include, for example, 1) peptides such as soluble peptides, including Ig-tailed fusion peptides and members of random peptide libraries (see, e.g., Lam *et al.*, *Nature 354*:82-84 (1991); Houghten *et al.*, *Nature 354*:84-86 (1991)) and combinatorial chemistry-derived molecular libraries made of D- and/or L- configuration amino acids; 2) phosphopeptides (e.g., members of random and partially degenerate, directed phosphopeptide libraries, see, e.g., Songyang *et al.*, *Cell 72*:767-778 (1993)); 3) antibodies (e.g., polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, and single chain antibodies as well as Fab, $F(ab')_2$, Fab expression library fragments, and epitope-binding fragments of antibodies); and 4) small organic and inorganic molecules (e.g., molecules obtained from combinatorial and natural product libraries).

One candidate compound is a soluble fragment of the receptor that competes for ligand binding. Other candidate compounds include mutant transporters or appropriate fragments containing mutations that affect transporter function and thus compete for ligand. Accordingly, a fragment that competes for ligand, for example with a higher affinity, or a fragment that binds ligand but does not allow release, is encompassed by the invention.

The invention further includes other end point assays to identify compounds that modulate (stimulate or inhibit) transporter activity. The assays typically involve an assay of events in the signal transduction pathway that indicate transporter activity. Thus, the transport of a ligand, change in cell membrane potential, activation of a protein, a change in the expression of genes that are up- or down-regulated in response to the transporter protein dependent signal cascade can be assayed.

Any of the biological or biochemical functions mediated by the transporter can be used as an endpoint assay. These include all of the biochemical or biochemical/biological events described herein, in the references cited herein, incorporated by reference for these endpoint assay targets, and other functions known to those of ordinary skill in the art or that can be readily identified using the information provided in the Figures, particularly Figure 2. Specifically, a biological function of a cell or tissues that expresses the transporter can be assayed. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

Binding and/or activating compounds can also be screened by using chimeric transporter proteins in which the amino terminal extracellular domain, or parts thereof, the entire transmembrane domain or subregions, such as any of the seven transmembrane segments or any of the intracellular or extracellular loops and the carboxy terminal intracellular domain, or parts thereof, can be replaced by heterologous domains or subregions. For example, a ligand-binding region can be used that interacts with a different ligand then that which is recognized by the native transporter. Accordingly, a different set of signal transduction components is available as an end-point assay for activation. This

allows for assays to be performed in other than the specific host cell from which the transporter is derived.

The proteins of the present invention are also useful in competition binding assays in methods designed to discover compounds that interact with the transporter (e.g. binding partners and/or ligands). Thus, a compound is exposed to a transporter polypeptide under conditions that allow the compound to bind or to otherwise interact with the polypeptide. Soluble transporter polypeptide is also added to the mixture. If the test compound interacts with the soluble transporter polypeptide, it decreases the amount of complex formed or activity from the transporter target. This type of assay is particularly useful in cases in which compounds are sought that interact with specific regions of the transporter. Thus, the soluble polypeptide that competes with the target transporter region is designed to contain peptide sequences corresponding to the region of interest.

To perform cell free drug screening assays, it is sometimes desirable to immobilize either the transporter protein, or fragment, or its target molecule to facilitate separation of complexes from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay.

Techniques for immobilizing proteins on matrices can be used in the drug screening assays. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be bound to a matrix. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the cell lysates (e.g., ^{35}S-labeled) and the candidate compound, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly, or in the supernatant after the complexes are dissociated. Alternatively, the complexes can be dissociated from the matrix, separated by SDS-PAGE, and the level of transporter-binding protein found in the bead fraction quantitated from the gel using standard electrophoretic techniques. For example, either the polypeptide or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin using techniques well known in the art. Alternatively, antibodies reactive with the protein but which do not interfere with binding of the protein to its target

molecule can be derivatized to the wells of the plate, and the protein trapped in the wells by antibody conjugation. Preparations of a transporter-binding protein and a candidate compound are incubated in the transporter protein-presenting wells and the amount of complex trapped in the well can be quantitated. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the transporter protein target molecule, or which are reactive with transporter protein and compete with the target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule.

Agents that modulate one of the transporters of the present invention can be identified using one or more of the above assays, alone or in combination. It is generally preferable to use a cell-based or cell free system first and then confirm activity in an animal or other model system. Such model systems are well known in the art and can readily be employed in this context.

Modulators of transporter protein activity identified according to these drug screening assays can be used to treat a subject with a disorder mediated by the transporter pathway, by treating cells or tissues that express the transporter. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. These methods of treatment include the steps of administering a modulator of transporter activity in a pharmaceutical composition to a subject in need of such treatment, the modulator being identified as described herein.

In yet another aspect of the invention, the transporter proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with the transporter and are involved in transporter activity. Such transporter-binding proteins are also likely to be involved in the propagation of signals by the transporter proteins or transporter targets as, for example, downstream elements of a transporter-mediated signaling pathway. Alternatively, such transporter-binding proteins are likely to be transporter inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a transporter protein is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a transporter-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the transporter protein.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a transporter-modulating agent, an antisense transporter nucleic acid molecule, a transporter-specific antibody, or a transporter-binding partner) can be used in an animal or other model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal or other model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

The transporter proteins of the present invention are also useful to provide a target for diagnosing a disease or predisposition to disease mediated by the peptide. Accordingly, the invention provides methods for detecting the presence, or levels of, the protein (or encoding mRNA) in a cell, tissue, or organism. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. The method involves contacting a biological sample with a compound capable of interacting with the transporter protein such that the interaction can be detected. Such an

assay can be provided in a single detection format or a multi-detection format such as an antibody chip array.

One agent for detecting a protein in a sample is an antibody capable of selectively binding to protein. A biological sample includes tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject.

The peptides of the present invention also provide targets for diagnosing active protein activity, disease, or predisposition to disease, in a patient having a variant peptide, particularly activities and conditions that are known for other members of the family of proteins to which the present one belongs. Thus, the peptide can be isolated from a biological sample and assayed for the presence of a genetic mutation that results in aberrant peptide. This includes amino acid substitution, deletion, insertion, rearrangement, (as the result of aberrant splicing events), and inappropriate post-translational modification. Analytic methods include altered electrophoretic mobility, altered tryptic peptide digest, altered transporter activity in cell-based or cell-free assay, alteration in ligand or antibody-binding pattern, altered isoelectric point, direct amino acid sequencing, and any other of the known assay techniques useful for detecting mutations in a protein. Such an assay can be provided in a single detection format or a multi-detection format such as an antibody chip array.

In vitro techniques for detection of peptide include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence using a detection reagent, such as an antibody or protein binding agent. Alternatively, the peptide can be detected *in vivo* in a subject by introducing into the subject a labeled anti-peptide antibody or other types of detection agent. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques. Particularly useful are methods that detect the allelic variant of a peptide expressed in a subject and methods which detect fragments of a peptide in a sample.

The peptides are also useful in pharmacogenomic analysis. Pharmacogenomics deal with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, e.g., Eichelbaum, M. (*Clin. Exp. Pharmacol. Physiol.* 23(10-11):983-985 (1996)), and Linder, M.W. (*Clin. Chem.* 43(2):254-266 (1997)). The clinical outcomes of these variations result in severe toxicity of

therapeutic drugs in certain individuals or therapeutic failure of drugs in certain individuals as a result of individual variation in metabolism. Thus, the genotype of the individual can determine the way a therapeutic compound acts on the body or the way the body metabolizes the compound. Further, the activity of drug metabolizing enzymes effects both the intensity and duration of drug action. Thus, the pharmacogenomics of the individual permit the selection of effective compounds and effective dosages of such compounds for prophylactic or therapeutic treatment based on the individual's genotype. The discovery of genetic polymorphisms in some drug metabolizing enzymes has explained why some patients do not obtain the expected drug effects, show an exaggerated drug effect, or experience serious toxicity from standard drug dosages. Polymorphisms can be expressed in the phenotype of the extensive metabolizer and the phenotype of the poor metabolizer. Accordingly, genetic polymorphism may lead to allelic protein variants of the transporter protein in which one or more of the transporter functions in one population is different from those in another population. The peptides thus allow a target to ascertain a genetic predisposition that can affect treatment modality. Thus, in a ligand-based treatment, polymorphism may give rise to amino terminal extracellular domains and/or other ligand-binding regions that are more or less active in ligand binding, and transporter activation. Accordingly, ligand dosage would necessarily be modified to maximize the therapeutic effect within a given population containing a polymorphism. As an alternative to genotyping, specific polymorphic peptides could be identified.

The peptides are also useful for treating a disorder characterized by an absence of, inappropriate, or unwanted expression of the protein. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. Accordingly, methods for treatment include the use of the transporter protein or fragments.

Antibodies

The invention also provides antibodies that selectively bind to one of the peptides of the present invention, a protein comprising such a peptide, as well as variants and fragments thereof. As used herein, an antibody selectively binds a target peptide when it binds the target peptide and does not significantly bind to unrelated proteins. An antibody is still

considered to selectively bind a peptide even if it also binds to other proteins that are not substantially homologous with the target peptide so long as such proteins share homology with a fragment or domain of the peptide target of the antibody. In this case, it would be understood that antibody binding to the peptide is still selective despite some degree of cross-reactivity.

As used herein, an antibody is defined in terms consistent with that recognized within the art: they are multi-subunit proteins produced by a mammalian organism in response to an antigen challenge. The antibodies of the present invention include polyclonal antibodies and monoclonal antibodies, as well as fragments of such antibodies, including, but not limited to, Fab or $F(ab')_2$, and Fv fragments.

Many methods are known for generating and/or identifying antibodies to a given target peptide. Several such methods are described by Harlow, Antibodies, Cold Spring Harbor Press, (1989).

In general, to generate antibodies, an isolated peptide is used as an immunogen and is administered to a mammalian organism, such as a rat, rabbit or mouse. The full-length protein, an antigenic peptide fragment or a fusion protein can be used. Particularly important fragments are those covering functional domains, such as the domains identified in Figure 2, and domain of sequence homology or divergence amongst the family, such as those that can readily be identified using protein alignment methods and as presented in the Figures.

Antibodies are preferably prepared from regions or discrete fragments of the transporter proteins. Antibodies can be prepared from any region of the peptide as described herein. However, preferred regions will include those involved in function/activity and/or transporter/binding partner interaction. Figure 2 can be used to identify particularly important regions while sequence alignment can be used to identify conserved and unique sequence fragments.

An antigenic fragment will typically comprise at least 8 contiguous amino acid residues. The antigenic peptide can comprise, however, at least 10, 12, 14, 16 or more amino acid residues. Such fragments can be selected on a physical property, such as fragments correspond to regions that are located on the surface of the protein, e.g., hydrophilic regions or can be selected based on sequence uniqueness (see Figure 2).

Detection on an antibody of the present invention can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

Antibody Uses

The antibodies can be used to isolate one of the proteins of the present invention by standard techniques, such as affinity chromatography or immunoprecipitation. The antibodies can facilitate the purification of the natural protein from cells and recombinantly produced protein expressed in host cells. In addition, such antibodies are useful to detect the presence of one of the proteins of the present invention in cells or tissues to determine the pattern of expression of the protein among various tissues in an organism and over the course of normal development. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. Further, such antibodies can be used to detect protein *in situ*, *in vitro*, or in a cell lysate or supernatant in order to evaluate the abundance and pattern of expression. Also, such antibodies can be used to assess abnormal tissue distribution or abnormal expression during development or progression of a biological condition. Antibody detection of circulating fragments of the full length protein can be used to identify turnover.

Further, the antibodies can be used to assess expression in disease states such as in active stages of the disease or in an individual with a predisposition toward disease related to the protein's function. When a disorder is caused by an inappropriate tissue distribution, developmental expression, level of expression of the protein, or expressed/processed form, the antibody can be prepared against the normal protein. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. If a disorder is characterized by a specific mutation in the protein, antibodies specific for this mutant protein can be used to assay for the presence of the specific mutant protein.

The antibodies can also be used to assess normal and aberrant subcellular localization of cells in the various tissues in an organism. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. The diagnostic uses can be applied, not only in genetic testing, but also in monitoring a treatment modality. Accordingly, where treatment is ultimately aimed at correcting expression level or the presence of aberrant sequence and aberrant tissue distribution or developmental expression, antibodies directed against the protein or relevant fragments can be used to monitor therapeutic efficacy.

Additionally, antibodies are useful in pharmacogenomic analysis. Thus, antibodies prepared against polymorphic proteins can be used to identify individuals that require modified treatment modalities. The antibodies are also useful as diagnostic tools as an immunological marker for aberrant protein analyzed by electrophoretic mobility, isoelectric point, tryptic peptide digest, and other physical assays known to those in the art.

The antibodies are also useful for tissue typing. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. Thus, where a specific protein has been correlated with expression in a specific tissue, antibodies that are specific for this protein can be used to identify a tissue type.

The antibodies are also useful for inhibiting protein function, for example, blocking the binding of the transporter peptide to a binding partner such as a ligand or protein binding partner. These uses can also be applied in a therapeutic context in which treatment involves inhibiting the protein's function. An antibody can be used, for example, to block binding,

thus modulating (agonizing or antagonizing) the peptides activity. Antibodies can be prepared against specific fragments containing sites required for function or against intact protein that is associated with a cell or cell membrane. See Figure 2 for structural information relating to the proteins of the present invention.

The invention also encompasses kits for using antibodies to detect the presence of a protein in a biological sample. The kit can comprise antibodies such as a labeled or labelable antibody and a compound or agent for detecting protein in a biological sample; means for determining the amount of protein in the sample; means for comparing the amount of protein in the sample with a standard; and instructions for use. Such a kit can be supplied to detect a single protein or epitope or can be configured to detect one of a multitude of epitopes, such as in an antibody detection array. Arrays are described in detail below for nucleic acid arrays and similar methods have been developed for antibody arrays.

Nucleic Acid Molecules

The present invention further provides isolated nucleic acid molecules that encode a transporter peptide or protein of the present invention (cDNA, transcript and genomic sequence). Such nucleic acid molecules will consist of, consist essentially of, or comprise a nucleotide sequence that encodes one of the transporter peptides of the present invention, an allelic variant thereof, or an ortholog or paralog thereof.

As used herein, an "isolated" nucleic acid molecule is one that is separated from other nucleic acid present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences that naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. However, there can be some flanking nucleotide sequences, for example up to about 5KB, 4KB, 3KB, 2KB, or 1KB or less, particularly contiguous peptide encoding sequences and peptide encoding sequences within the same gene but separated by introns in the genomic sequence. The important point is that the nucleic acid is isolated from remote and unimportant flanking sequences such that it can be subjected to the specific manipulations described herein such as recombinant expression, preparation of probes and primers, and other uses specific to the nucleic acid sequences.

Moreover, an "isolated" nucleic acid molecule, such as a transcript/cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. However, the nucleic acid molecule can be fused to other coding or regulatory sequences and still be considered isolated.

For example, recombinant DNA molecules contained in a vector are considered isolated. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the isolated DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

Accordingly, the present invention provides nucleic acid molecules that consist of the nucleotide sequence shown in Figure 1 or 3 (SEQ ID NO:1, transcript sequence and SEQ ID NO:3, genomic sequence), or any nucleic acid molecule that encodes the protein provided in Figure 2, SEQ ID NO:2. A nucleic acid molecule consists of a nucleotide sequence when the nucleotide sequence is the complete nucleotide sequence of the nucleic acid molecule.

The present invention further provides nucleic acid molecules that consist essentially of the nucleotide sequence shown in Figure 1 or 3 (SEQ ID NO:1, transcript sequence and SEQ ID NO:3, genomic sequence), or any nucleic acid molecule that encodes the protein provided in Figure 2, SEQ ID NO:2. A nucleic acid molecule consists essentially of a nucleotide sequence when such a nucleotide sequence is present with only a few additional nucleic acid residues in the final nucleic acid molecule.

The present invention further provides nucleic acid molecules that comprise the nucleotide sequences shown in Figure 1 or 3 (SEQ ID NO:1, transcript sequence and SEQ ID NO:3, genomic sequence), or any nucleic acid molecule that encodes the protein provided in Figure 2, SEQ ID NO:2. A nucleic acid molecule comprises a nucleotide sequence when the nucleotide sequence is at least part of the final nucleotide sequence of the nucleic acid molecule. In such a fashion, the nucleic acid molecule can be only the nucleotide sequence or have additional nucleic acid residues, such as nucleic acid residues that are naturally associated with it or heterologous nucleotide sequences. Such a nucleic

acid molecule can have a few additional nucleotides or can comprise several hundred or more additional nucleotides. A brief description of how various types of these nucleic acid molecules can be readily made/isolated is provided below.

In Figures 1 and 3, both coding and non-coding sequences are provided. Because of the source of the present invention, humans genomic sequence (Figure 3) and cDNA/transcript sequences (Figure 1), the nucleic acid molecules in the Figures will contain genomic intronic sequences, 5' and 3' non-coding sequences, gene regulatory regions and non-coding intergenic sequences. In general such sequence features are either noted in Figures 1 and 3 or can readily be identified using computational tools known in the art. As discussed below, some of the non-coding regions, particularly gene regulatory elements such as promoters, are useful for a variety of purposes, e.g. control of heterologous gene expression, target for identifying gene activity modulating compounds, and are particularly claimed as fragments of the genomic sequence provided herein.

The isolated nucleic acid molecules can encode the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature peptide (when the mature form has more than one peptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, facilitate protein trafficking, prolong or shorten protein half-life or facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in situ*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

As mentioned above, the isolated nucleic acid molecules include, but are not limited to, the sequence encoding the transporter peptide alone, the sequence encoding the mature peptide and additional coding sequences, such as a leader or secretory sequence (e.g., a pre-pro or pro-protein sequence), the sequence encoding the mature peptide, with or without the additional coding sequences, plus additional non-coding sequences, for example introns and non-coding 5' and 3' sequences such as transcribed but non-translated sequences that play a role in transcription, mRNA processing (including splicing and polyadenylation signals), ribosome binding and stability of mRNA. In addition, the nucleic acid molecule may be fused to a marker sequence encoding, for example, a peptide that facilitates purification.

Isolated nucleic acid molecules can be in the form of RNA, such as mRNA, or in the form DNA, including cDNA and genomic DNA obtained by cloning or produced by

chemical synthetic techniques or by a combination thereof. The nucleic acid, especially DNA, can be double-stranded or single-stranded. Single-stranded nucleic acid can be the coding strand (sense strand) or the non-coding strand (anti-sense strand).

The invention further provides nucleic acid molecules that encode fragments of the peptides of the present invention as well as nucleic acid molecules that encode obvious variants of the transporter proteins of the present invention that are described above. Such nucleic acid molecules may be naturally occurring, such as allelic variants (same locus), paralogs (different locus), and orthologs (different organism), or may be constructed by recombinant DNA methods or by chemical synthesis. Such non-naturally occurring variants may be made by mutagenesis techniques, including those applied to nucleic acid molecules, cells, or organisms. Accordingly, as discussed above, the variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions.

The present invention further provides non-coding fragments of the nucleic acid molecules provided in Figures 1 and 3. Preferred non-coding fragments include, but are not limited to, promoter sequences, enhancer sequences, gene modulating sequences and gene termination sequences. Such fragments are useful in controlling heterologous gene expression and in developing screens to identify gene-modulating agents. A promoter can readily be identified as being 5' to the ATG start site in the genomic sequence provided in Figure 3.

A fragment comprises a contiguous nucleotide sequence greater than 12 or more nucleotides. Further, a fragment could at least 30, 40, 50, 100, 250 or 500 nucleotides in length. The length of the fragment will be based on its intended use. For example, the fragment can encode epitope bearing regions of the peptide, or can be useful as DNA probes and primers. Such fragments can be isolated using the known nucleotide sequence to synthesize an oligonucleotide probe. A labeled probe can then be used to screen a cDNA library, genomic DNA library, or mRNA to isolate nucleic acid corresponding to the coding region. Further, primers can be used in PCR reactions to clone specific regions of gene.

A probe/primer typically comprises substantially a purified oligonucleotide or oligonucleotide pair. The oligonucleotide typically comprises a region of nucleotide

sequence that hybridizes under stringent conditions to at least about 12, 20, 25, 40, 50 or more consecutive nucleotides.

Orthologs, homologs, and allelic variants can be identified using methods well known in the art. As described in the Peptide Section, these variants comprise a nucleotide sequence encoding a peptide that is typically 60-70%, 70-80%, 80-90%, and more typically at least about 90-95% or more homologous to the nucleotide sequence shown in the Figure sheets or a fragment of this sequence. Such nucleic acid molecules can readily be identified as being able to hybridize under moderate to stringent conditions, to the nucleotide sequence shown in the Figure sheets or a fragment of the sequence. Allelic variants can readily be determined by genetic locus of the encoding gene. As indicated by the data presented in Figure 3, the map position was determined to be on chromosome 14 by ePCR.

Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 140 SNP variants were found, including 6 indels (indicated by a "-") and 1 SNPs in exons. The others were found in in introns and regions 5' and 3' of the ORF. Such SNPs in introns and outside the ORF may affect control/regulatory elements.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences encoding a peptide at least 60-70% homologous to each other typically remain hybridized to each other. The conditions can be such that sequences at least about 60%, at least about 70%, or at least about 80% or more homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. One example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65C. Examples of moderate to low stringency hybridization conditions are well known in the art.

Nucleic Acid Molecule Uses

The nucleic acid molecules of the present invention are useful for probes, primers, chemical intermediates, and in biological assays. The nucleic acid molecules are useful as a

hybridization probe for messenger RNA, transcript/cDNA and genomic DNA to isolate full-length cDNA and genomic clones encoding the peptide described in Figure 2 and to isolate cDNA and genomic clones that correspond to variants (alleles, orthologs, etc.) producing the same or related peptides shown in Figure 2. 140 SNPs, including 6 indels, have been identified in the gene encoding the transporter protein provided by the present invention and are given in Figure 3.

The probe can correspond to any sequence along the entire length of the nucleic acid molecules provided in the Figures. Accordingly, it could be derived from 5' noncoding regions, the coding region, and 3' noncoding regions. However, as discussed, fragments are not to be construed as encompassing fragments disclosed prior to the present invention.

The nucleic acid molecules are also useful as primers for PCR to amplify any given region of a nucleic acid molecule and are useful to synthesize antisense molecules of desired length and sequence.

The nucleic acid molecules are also useful for constructing recombinant vectors. Such vectors include expression vectors that express a portion of, or all of, the peptide sequences. Vectors also include insertion vectors, used to integrate into another nucleic acid molecule sequence, such as into the cellular genome, to alter *in situ* expression of a gene and/or gene product. For example, an endogenous coding sequence can be replaced via homologous recombination with all or part of the coding region containing one or more specifically introduced mutations.

The nucleic acid molecules are also useful for expressing antigenic portions of the proteins.

The nucleic acid molecules are also useful as probes for determining the chromosomal positions of the nucleic acid molecules by means of *in situ* hybridization methods. As indicated by the data presented in Figure 3, the map position was determined to be on chromosome 14 by ePCR.

The nucleic acid molecules are also useful in making vectors containing the gene regulatory regions of the nucleic acid molecules of the present invention.

The nucleic acid molecules are also useful for designing ribozymes corresponding to all, or a part, of the mRNA produced from the nucleic acid molecules described herein.

The nucleic acid molecules are also useful for making vectors that express part, or all, of the peptides.

The nucleic acid molecules are also useful for constructing host cells expressing a part, or all, of the nucleic acid molecules and peptides.

The nucleic acid molecules are also useful for constructing transgenic animals expressing all, or a part, of the nucleic acid molecules and peptides.

The nucleic acid molecules are also useful as hybridization probes for determining the presence, level, form and distribution of nucleic acid expression. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

Accordingly, the probes can be used to detect the presence of, or to determine levels of, a specific nucleic acid molecule in cells, tissues, and in organisms. The nucleic acid whose level is determined can be DNA or RNA. Accordingly, probes corresponding to the peptides described herein can be used to assess expression and/or gene copy number in a given cell, tissue, or organism. These uses are relevant for diagnosis of disorders involving an increase or decrease in transporter protein expression relative to normal results.

In vitro techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detecting DNA include Southern hybridizations and *in situ* hybridization.

Probes can be used as a part of a diagnostic test kit for identifying cells or tissues that express a transporter protein, such as by measuring a level of a transporter-encoding nucleic acid in a sample of cells from a subject e.g., mRNA or genomic DNA, or determining if a transporter gene has been mutated. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

Nucleic acid expression assays are useful for drug screening to identify compounds that modulate transporter nucleic acid expression.

The invention thus provides a method for identifying a compound that can be used to treat a disorder associated with nucleic acid expression of the transporter gene, particularly biological and pathological processes that are mediated by the transporter in cells and tissues that express it. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. The method typically includes assaying the ability of the compound to modulate the expression of the transporter nucleic acid and thus identifying a compound that can be used to treat a disorder characterized by undesired transporter nucleic acid expression. The assays can be performed in cell-based and cell-free systems. Cell-based assays include cells naturally expressing the transporter nucleic acid or recombinant cells genetically engineered to express specific nucleic acid sequences.

The assay for transporter nucleic acid expression can involve direct assay of nucleic acid levels, such as mRNA levels, or on collateral compounds involved in the signal pathway. Further, the expression of genes that are up- or down-regulated in response to the transporter protein signal pathway can also be assayed. In this embodiment the regulatory regions of these genes can be operably linked to a reporter gene such as luciferase.

Thus, modulators of transporter gene expression can be identified in a method wherein a cell is contacted with a candidate compound and the expression of mRNA determined. The level of expression of transporter mRNA in the presence of the candidate compound is compared to the level of expression of transporter mRNA in the absence of the candidate compound. The candidate compound can then be identified as a modulator of nucleic acid expression based on this comparison and be used, for example to treat a disorder characterized by aberrant nucleic acid expression. When expression of mRNA is statistically significantly greater in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of nucleic acid expression. When nucleic acid expression is statistically significantly less in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of nucleic acid expression.

The invention further provides methods of treatment, with the nucleic acid as a target, using a compound identified through drug screening as a gene modulator to modulate transporter nucleic acid expression in cells and tissues that express the transporter. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. Modulation includes both up-regulation (i.e. activation or agonization) or down-regulation (suppression or antagonization) or nucleic acid expression.

Alternatively, a modulator for transporter nucleic acid expression can be a small molecule or drug identified using the screening assays described herein as long as the drug or small molecule inhibits the transporter nucleic acid expression in the cells and tissues that express the protein. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

The nucleic acid molecules are also useful for monitoring the effectiveness of modulating compounds on the expression or activity of the transporter gene in clinical trials or in a treatment regimen. Thus, the gene expression pattern can serve as a barometer for the continuing effectiveness of treatment with the compound, particularly with compounds to which a patient can develop resistance. The gene expression pattern can also serve as a marker indicative of a physiological response of the affected cells to the compound. Accordingly, such monitoring would allow either increased administration of the compound or the administration of alternative compounds to which the patient has not become resistant. Similarly, if the level of nucleic acid expression falls below a desirable level, administration of the compound could be commensurately decreased.

The nucleic acid molecules are also useful in diagnostic assays for qualitative changes in transporter nucleic acid expression, and particularly in qualitative changes that lead to pathology. The nucleic acid molecules can be used to detect mutations in transporter genes and gene expression products such as mRNA. The nucleic acid molecules can be used as hybridization probes to detect naturally occurring genetic mutations in the transporter gene and thereby to determine whether a subject with the mutation is at risk for a

disorder caused by the mutation. Mutations include deletion, addition, or substitution of one or more nucleotides in the gene, chromosomal rearrangement, such as inversion or transposition, modification of genomic DNA, such as aberrant methylation patterns or changes in gene copy number, such as amplification. Detection of a mutated form of the transporter gene associated with a dysfunction provides a diagnostic tool for an active disease or susceptibility to disease when the disease results from overexpression, underexpression, or altered expression of a transporter protein.

Individuals carrying mutations in the transporter gene can be detected at the nucleic acid level by a variety of techniques. Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 140 SNP variants were found, including 6 indels (indicated by a "-") and 1 SNPs in exons. The others were found in in introns and regions 5' and 3' of the ORF. Such SNPs in introns and outside the ORF may affect control/regulatory elements. As indicated by the data presented in Figure 3, the map position was determined to be on chromosome 14 by ePCR. Genomic DNA can be analyzed directly or can be amplified by using PCR prior to analysis. RNA or cDNA can be used in the same way. In some uses, detection of the mutation involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran *et al.*, *Science 241*:1077-1080 (1988); and Nakazawa *et al.*, *PNAS 91*:360-364 (1994)), the latter of which can be particularly useful for detecting point mutations in the gene (see Abravaya *et al.*, *Nucleic Acids Res. 23*:675-682 (1995)). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a gene under conditions such that hybridization and amplification of the gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. Deletions and insertions can be detected by a change in size of the amplified product compared to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to normal RNA or antisense DNA sequences.

Alternatively, mutations in a transporter gene can be directly identified, for example, by alterations in restriction enzyme digestion patterns determined by gel electrophoresis.

Further, sequence-specific ribozymes (U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site. Perfectly matched sequences can be distinguished from mismatched sequences by nuclease cleavage digestion assays or by differences in melting temperature.

Sequence changes at specific locations can also be assessed by nuclease protection assays such as RNase and S1 protection or the chemical cleavage method. Furthermore, sequence differences between a mutant transporter gene and a wild-type gene can be determined by direct DNA sequencing. A variety of automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve, C.W., (1995) *Biotechniques 19*:448), including sequencing by mass spectrometry (see, e.g., PCT International Publication No. WO 94/16101; Cohen *et al.*, *Adv. Chromatogr. 36*:127-162 (1996); and Griffin *et al.*, *Appl. Biochem. Biotechnol. 38*:147-159 (1993)).

Other methods for detecting mutations in the gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA duplexes (Myers *et al.*, *Science 230*:1242 (1985)); Cotton *et al.*, *PNAS 85*:4397 (1988); Saleeba *et al.*, *Meth. Enzymol. 217*:286-295 (1992)), electrophoretic mobility of mutant and wild type nucleic acid is compared (Orita *et al.*, *PNAS 86*:2766 (1989); Cotton *et al.*, *Mutat. Res. 285*:125-144 (1993); and Hayashi *et al.*, *Genet. Anal. Tech. Appl. 9*:73-79 (1992)), and movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (Myers *et al.*, *Nature 313*:495 (1985)). Examples of other techniques for detecting point mutations include selective oligonucleotide hybridization, selective amplification, and selective primer extension.

The nucleic acid molecules are also useful for testing an individual for a genotype that while not necessarily causing the disease, nevertheless affects the treatment modality. Thus, the nucleic acid molecules can be used to study the relationship between an individual's genotype and the individual's response to a compound used for treatment (pharmacogenomic relationship). Accordingly, the nucleic acid molecules described herein can be used to assess the mutation content of the transporter gene in an individual in order to

select an appropriate compound or dosage regimen for treatment. Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 140 SNP variants were found, including 6 indels (indicated by a "-") and 1 SNPs in exons. The others were found in in introns and regions 5' and 3' of the ORF. Such SNPs in introns and outside the ORF may affect control/regulatory elements.

Thus nucleic acid molecules displaying genetic variations that affect treatment provide a diagnostic target that can be used to tailor treatment in an individual. Accordingly, the production of recombinant cells and animals containing these polymorphisms allow effective clinical design of treatment compounds and dosage regimens.

The nucleic acid molecules are thus useful as antisense constructs to control transporter gene expression in cells, tissues, and organisms. A DNA antisense nucleic acid molecule is designed to be complementary to a region of the gene involved in transcription, preventing transcription and hence production of transporter protein. An antisense RNA or DNA nucleic acid molecule would hybridize to the mRNA and thus block translation of mRNA into transporter protein.

Alternatively, a class of antisense molecules can be used to inactivate mRNA in order to decrease expression of transporter nucleic acid. Accordingly, these molecules can treat a disorder characterized by abnormal or undesired transporter nucleic acid expression. This technique involves cleavage by means of ribozymes containing nucleotide sequences complementary to one or more regions in the mRNA that attenuate the ability of the mRNA to be translated. Possible regions include coding regions and particularly coding regions corresponding to the catalytic and other functional activities of the transporter protein, such as ligand binding.

The nucleic acid molecules also provide vectors for gene therapy in patients containing cells that are aberrant in transporter gene expression. Thus, recombinant cells, which include the patient's cells that have been engineered *ex vivo* and returned to the patient, are introduced into an individual where the cells produce the desired transporter protein to treat the individual.

The invention also encompasses kits for detecting the presence of a transporter nucleic acid in a biological sample. Experimental data as provided in Figure 1 indicates that

sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. For example, the kit can comprise reagents such as a labeled or labelable nucleic acid or agent capable of detecting transporter nucleic acid in a biological sample; means for determining the amount of transporter nucleic acid in the sample; and means for comparing the amount of transporter nucleic acid in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect transporter protein mRNA or DNA.

Nucleic Acid Arrays

The present invention further provides nucleic acid detection kits, such as arrays or microarrays of nucleic acid molecules that are based on the sequence information provided in Figures 1 and 3 (SEQ ID NOS:1 and 3).

As used herein "Arrays" or "Microarrays" refers to an array of distinct polynucleotides or oligonucleotides synthesized on a substrate, such as paper, nylon or other type of membrane, filter, chip, glass slide, or any other suitable solid support. In one embodiment, the microarray is prepared and used according to the methods described in US Patent 5,837,832, Chee *et al.*, PCT application W095/11995 (Chee *et al.*), Lockhart, D. J. *et al.* (1996; Nat. Biotech. 14: 1675-1680) and Schena, M. *et al.* (1996; Proc. Natl. Acad. Sci. 93: 10614-10619), all of which are incorporated herein in their entirety by reference. In other embodiments, such arrays are produced by the methods described by Brown *et al.*, US Patent No. 5,807,522.

The microarray or detection kit is preferably composed of a large number of unique, single-stranded nucleic acid sequences, usually either synthetic antisense oligonucleotides or fragments of cDNAs, fixed to a solid support. The oligonucleotides are preferably about 6-60 nucleotides in length, more preferably 15-30 nucleotides in length, and most preferably about 20-25 nucleotides in length. For a certain type of microarray or detection kit, it may be preferable to use oligonucleotides that are only 7-20 nucleotides in length. The microarray or detection kit may contain oligonucleotides

that cover the known 5', or 3', sequence, sequential oligonucleotides that cover the full length sequence; or unique oligonucleotides selected from particular areas along the length of the sequence. Polynucleotides used in the microarray or detection kit may be oligonucleotides that are specific to a gene or genes of interest.

In order to produce oligonucleotides to a known sequence for a microarray or detection kit, the gene(s) of interest (or an ORF identified from the contigs of the present invention) is typically examined using a computer algorithm which starts at the 5' or at the 3' end of the nucleotide sequence. Typical algorithms will then identify oligomers of defined length that are unique to the gene, have a GC content within a range suitable for hybridization, and lack predicted secondary structure that may interfere with hybridization. In certain situations it may be appropriate to use pairs of oligonucleotides on a microarray or detection kit. The "pairs" will be identical, except for one nucleotide that preferably is located in the center of the sequence. The second oligonucleotide in the pair (mismatched by one) serves as a control. The number of oligonucleotide pairs may range from two to one million. The oligomers are synthesized at designated areas on a substrate using a light-directed chemical process. The substrate may be paper, nylon or other type of membrane, filter, chip, glass slide or any other suitable solid support.

In another aspect, an oligonucleotide may be synthesized on the surface of the substrate by using a chemical coupling procedure and an ink jet application apparatus, as described in PCT application W095/251116 (Baldeschweiler *et al.*) which is incorporated herein in its entirety by reference. In another aspect, a "gridded" array analogous to a dot (or slot) blot may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedures. An array, such as those described above, may be produced by hand or by using available devices (slot blot or dot blot apparatus), materials (any suitable solid support), and machines (including robotic instruments), and may contain 8, 24, 96, 384, 1536, 6144 or more oligonucleotides, or any other number between two and one million which lends itself to the efficient use of commercially available instrumentation.

In order to conduct sample analysis using a microarray or detection kit, the RNA or DNA from a biological sample is made into hybridization probes. The mRNA is isolated, and cDNA is produced and used as a template to make antisense RNA (aRNA).

The aRNA is amplified in the presence of fluorescent nucleotides, and labeled probes are incubated with the microarray or detection kit so that the probe sequences hybridize to complementary oligonucleotides of the microarray or detection kit. Incubation conditions are adjusted so that hybridization occurs with precise complementary matches or with various degrees of less complementarity. After removal of nonhybridized probes, a scanner is used to determine the levels and patterns of fluorescence. The scanned images are examined to determine degree of complementarity and the relative abundance of each oligonucleotide sequence on the microarray or detection kit. The biological samples may be obtained from any bodily fluids (such as blood, urine, saliva, phlegm, gastric juices, etc.), cultured cells, biopsies, or other tissue preparations. A detection system may be used to measure the absence, presence, and amount of hybridization for all of the distinct sequences simultaneously. This data may be used for large-scale correlation studies on the sequences, expression patterns, mutations, variants, or polymorphisms among samples.

Using such arrays, the present invention provides methods to identify the expression of the transporter proteins/peptides of the present invention. In detail, such methods comprise incubating a test sample with one or more nucleic acid molecules and assaying for binding of the nucleic acid molecule with components within the test sample. Such assays will typically involve arrays comprising many genes, at least one of which is a gene of the present invention and or alleles of the transporter gene of the present invention. Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 140 SNP variants were found, including 6 indels (indicated by a "-") and 1 SNPs in exons. The others were found in in introns and regions 5' and 3' of the ORF. Such SNPs in introns and outside the ORF may affect control/regulatory elements.

Conditions for incubating a nucleic acid molecule with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid molecule used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or array assay formats can readily be adapted to employ the novel fragments of the Human genome disclosed herein. Examples of such assays can be found

in Chard, T, *An Introduction to Radioimmunoassay and Related Techniques,* Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G. R. *et al., Techniques in Immunocytochemistry,* Academic Press, Orlando, FL Vol. 1 (1 982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., *Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology,* Elsevier Science Publishers, Amsterdam, The Netherlands (1985).

The test samples of the present invention include cells, protein or membrane extracts of cells. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing nucleic acid extracts or of cells are well known in the art and can be readily be adapted in order to obtain a sample that is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention.

Specifically, the invention provides a compartmentalized kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the nucleic acid molecules that can bind to a fragment of the Human genome disclosed herein; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound nucleic acid.

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers, strips of plastic, glass or paper, or arraying material such as silica. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the nucleic acid probe, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound probe. One skilled in the art will readily recognize that the previously unidentified transporter gene of the present invention can be routinely identified using the sequence information disclosed herein can be readily incorporated

into one of the established kit formats which are well known in the art, particularly expression arrays.

Vectors/host cells

The invention also provides vectors containing the nucleic acid molecules described herein. The term "vector" refers to a vehicle, preferably a nucleic acid molecule, which can transport the nucleic acid molecules. When the vector is a nucleic acid molecule, the nucleic acid molecules are covalently linked to the vector nucleic acid. With this aspect of the invention, the vector includes a plasmid, single or double stranded phage, a single or double stranded RNA or DNA viral vector, or artificial chromosome, such as a BAC, PAC, YAC, OR MAC.

A vector can be maintained in the host cell as an extrachromosomal element where it replicates and produces additional copies of the nucleic acid molecules. Alternatively, the vector may integrate into the host cell genome and produce additional copies of the nucleic acid molecules when the host cell replicates.

The invention provides vectors for the maintenance (cloning vectors) or vectors for expression (expression vectors) of the nucleic acid molecules. The vectors can function in procaryotic or eukaryotic cells or in both (shuttle vectors).

Expression vectors contain cis-acting regulatory regions that are operably linked in the vector to the nucleic acid molecules such that transcription of the nucleic acid molecules is allowed in a host cell. The nucleic acid molecules can be introduced into the host cell with a separate nucleic acid molecule capable of affecting transcription. Thus, the second nucleic acid molecule may provide a trans-acting factor interacting with the cis-regulatory control region to allow transcription of the nucleic acid molecules from the vector. Alternatively, a trans-acting factor may be supplied by the host cell. Finally, a trans-acting factor can be produced from the vector itself. It is understood, however, that in some embodiments, transcription and/or translation of the nucleic acid molecules can occur in a cell-free system.

The regulatory sequence to which the nucleic acid molecules described herein can be operably linked include promoters for directing mRNA transcription. These include, but are not limited to, the left promoter from bacteriophage λ, the lac, TRP, and TAC promoters

from *E. coli*, the early and late promoters from SV40, the CMV immediate early promoter, the adenovirus early and late promoters, and retrovirus long-terminal repeats.

In addition to control regions that promote transcription, expression vectors may also include regions that modulate transcription, such as repressor binding sites and enhancers. Examples include the SV40 enhancer, the cytomegalovirus immediate early enhancer, polyoma enhancer, adenovirus enhancers, and retrovirus LTR enhancers.

In addition to containing sites for transcription initiation and control, expression vectors can also contain sequences necessary for transcription termination and, in the transcribed region a ribosome binding site for translation. Other regulatory control elements for expression include initiation and termination codons as well as polyadenylation signals. The person of ordinary skill in the art would be aware of the numerous regulatory sequences that are useful in expression vectors. Such regulatory sequences are described, for example, in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual. 2nd. ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

A variety of expression vectors can be used to express a nucleic acid molecule. Such vectors include chromosomal, episomal, and virus-derived vectors, for example vectors derived from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast chromosomal elements, including yeast artificial chromosomes, from viruses such as baculoviruses, papovaviruses such as SV40, Vaccinia viruses, adenoviruses, poxviruses, pseudorabies viruses, and retroviruses. Vectors may also be derived from combinations of these sources such as those derived from plasmid and bacteriophage genetic elements, e.g. cosmids and phagemids. Appropriate cloning and expression vectors for prokaryotic and eukaryotic hosts are described in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual. 2nd. ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

The regulatory sequence may provide constitutive expression in one or more host cells (i.e. tissue specific) or may provide for inducible expression in one or more cell types such as by temperature, nutrient additive, or exogenous factor such as a hormone or other ligand. A variety of vectors providing for constitutive and inducible expression in prokaryotic and eukaryotic hosts are well known to those of ordinary skill in the art.

The nucleic acid molecules can be inserted into the vector nucleic acid by well-known methodology. Generally, the DNA sequence that will ultimately be expressed is

joined to an expression vector by cleaving the DNA sequence and the expression vector with one or more restriction enzymes and then ligating the fragments together. Procedures for restriction enzyme digestion and ligation are well known to those of ordinary skill in the art.

The vector containing the appropriate nucleic acid molecule can be introduced into an appropriate host cell for propagation or expression using well-known techniques. Bacterial cells include, but are not limited to, *E. coli*, *Streptomyces*, and *Salmonella typhimurium*. Eukaryotic cells include, but are not limited to, yeast, insect cells such as *Drosophila*, animal cells such as COS and CHO cells, and plant cells.

As described herein, it may be desirable to express the peptide as a fusion protein. Accordingly, the invention provides fusion vectors that allow for the production of the peptides. Fusion vectors can increase the expression of a recombinant protein, increase the solubility of the recombinant protein, and aid in the purification of the protein by acting for example as a ligand for affinity purification. A proteolytic cleavage site may be introduced at the junction of the fusion moiety so that the desired peptide can ultimately be separated from the fusion moiety. Proteolytic enzymes include, but are not limited to, factor Xa, thrombin, and enterotransporter. Typical fusion expression vectors include pGEX (Smith *et al.*, *Gene 67*:31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, *Gene 69*:301-315 (1988)) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology 185*:60-89 (1990)).

Recombinant protein expression can be maximized in host bacteria by providing a genetic background wherein the host cell has an impaired capacity to proteolytically cleave the recombinant protein. (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Alternatively, the sequence of the nucleic acid molecule of interest can be altered to provide preferential codon usage for a specific host cell, for example *E. coli*. (Wada *et al.*, *Nucleic Acids Res. 20*:2111-2118 (1992)).

The nucleic acid molecules can also be expressed by expression vectors that are operative in yeast. Examples of vectors for expression in yeast e.g., *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, *EMBO J. 6*:229-234 (1987)), pMFa (Kurjan *et al.*, *Cell 30*:933-943(1982)), pJRY88 (Schultz *et al.*, *Gene 54*:113-123 (1987)), and pYES2 (Invitrogen Corporation, San Diego, CA).

The nucleic acid molecules can also be expressed in insect cells using, for example, baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith *et al.*, *Mol. Cell Biol. 3*:2156-2165 (1983)) and the pVL series (Lucklow *et al.*, *Virology 170*:31-39 (1989)).

In certain embodiments of the invention, the nucleic acid molecules described herein are expressed in mammalian cells using mammalian expression vectors. Examples of mammalian expression vectors include pCDM8 (Seed, B. *Nature 329*:840(1987)) and pMT2PC (Kaufman *et al.*, *EMBO J. 6*:187-195 (1987)).

The expression vectors listed herein are provided by way of example only of the well-known vectors available to those of ordinary skill in the art that would be useful to express the nucleic acid molecules. The person of ordinary skill in the art would be aware of other vectors suitable for maintenance propagation or expression of the nucleic acid molecules described herein. These are found for example in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

The invention also encompasses vectors in which the nucleic acid sequences described herein are cloned into the vector in reverse orientation, but operably linked to a regulatory sequence that permits transcription of antisense RNA. Thus, an antisense transcript can be produced to all, or to a portion, of the nucleic acid molecule sequences described herein, including both coding and non-coding regions. Expression of this antisense RNA is subject to each of the parameters described above in relation to expression of the sense RNA (regulatory sequences, constitutive or inducible expression, tissue-specific expression).

The invention also relates to recombinant host cells containing the vectors described herein. Host cells therefore include prokaryotic cells, lower eukaryotic cells such as yeast,

other eukaryotic cells such as insect cells, and higher eukaryotic cells such as mammalian cells.

The recombinant host cells are prepared by introducing the vector constructs described herein into the cells by techniques readily available to the person of ordinary skill in the art. These include, but are not limited to, calcium phosphate transfection, DEAE-dextran-mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, lipofection, and other techniques such as those found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Host cells can contain more than one vector. Thus, different nucleotide sequences can be introduced on different vectors of the same cell. Similarly, the nucleic acid molecules can be introduced either alone or with other nucleic acid molecules that are not related to the nucleic acid molecules such as those providing trans-acting factors for expression vectors. When more than one vector is introduced into a cell, the vectors can be introduced independently, co-introduced or joined to the nucleic acid molecule vector.

In the case of bacteriophage and viral vectors, these can be introduced into cells as packaged or encapsulated virus by standard procedures for infection and transduction. Viral vectors can be replication-competent or replication-defective. In the case in which viral replication is defective, replication will occur in host cells providing functions that complement the defects.

Vectors generally include selectable markers that enable the selection of the subpopulation of cells that contain the recombinant vector constructs. The marker can be contained in the same vector that contains the nucleic acid molecules described herein or may be on a separate vector. Markers include tetracycline or ampicillin-resistance genes for prokaryotic host cells and dihydrofolate reductase or neomycin resistance for eukaryotic host cells. However, any marker that provides selection for a phenotypic trait will be effective.

While the mature proteins can be produced in bacteria, yeast, mammalian cells, and other cells under the control of the appropriate regulatory sequences, cell- free transcription and translation systems can also be used to produce these proteins using RNA derived from the DNA constructs described herein.

Where secretion of the peptide is desired, which is difficult to achieve with multi-transmembrane domain containing proteins such as transporters, appropriate secretion signals are incorporated into the vector. The signal sequence can be endogenous to the peptides or heterologous to these peptides.

Where the peptide is not secreted into the medium, which is typically the case with transporters, the protein can be isolated from the host cell by standard disruption procedures, including freeze thaw, sonication, mechanical disruption, use of lysing agents and the like. The peptide can then be recovered and purified by well-known purification methods including ammonium sulfate precipitation, acid extraction, anion or cationic exchange chromatography, phosphocellulose chromatography, hydrophobic-interaction chromatography, affinity chromatography, hydroxylapatite chromatography, lectin chromatography, or high performance liquid chromatography.

It is also understood that depending upon the host cell in recombinant production of the peptides described herein, the peptides can have various glycosylation patterns, depending upon the cell, or maybe non-glycosylated as when produced in bacteria. In addition, the peptides may include an initial modified methionine in some cases as a result of a host-mediated process.

Uses of vectors and host cells

The recombinant host cells expressing the peptides described herein have a variety of uses. First, the cells are useful for producing a transporter protein or peptide that can be further purified to produce desired amounts of transporter protein or fragments. Thus, host cells containing expression vectors are useful for peptide production.

Host cells are also useful for conducting cell-based assays involving the transporter protein or transporter protein fragments, such as those described above as well as other formats known in the art. Thus, a recombinant host cell expressing a native transporter protein is useful for assaying compounds that stimulate or inhibit transporter protein function.

Host cells are also useful for identifying transporter protein mutants in which these functions are affected. If the mutants naturally occur and give rise to a pathology, host cells containing the mutations are useful to assay compounds that have a desired effect on the

mutant transporter protein (for example, stimulating or inhibiting function) which may not be indicated by their effect on the native transporter protein.

Genetically engineered host cells can be further used to produce non-human transgenic animals. A transgenic animal is preferably a mammal, for example a rodent, such as a rat or mouse, in which one or more of the cells of the animal include a transgene. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal in one or more cell types or tissues of the transgenic animal. These animals are useful for studying the function of a transporter protein and identifying and evaluating modulators of transporter protein activity. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, and amphibians.

A transgenic animal can be produced by introducing nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Any of the transporter protein nucleotide sequences can be introduced as a transgene into the genome of a non-human animal, such as a mouse.

Any of the regulatory or other sequences useful in expression vectors can form part of the transgenic sequence. This includes intronic sequences and polyadenylation signals, if not already included. A tissue-specific regulatory sequence(s) can be operably linked to the transgene to direct expression of the transporter protein to particular cells.

Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the transgene in its genome and/or expression of transgenic mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene can further be bred to other transgenic animals carrying other transgenes. A transgenic animal also includes animals in which the entire animal or tissues

in the animal have been produced using the homologously recombinant host cells described herein.

In another embodiment, transgenic non-human animals can be produced which contain selected systems that allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, e.g., Lakso *et al. PNAS 89*:6232-6236 (1992). Another example of a recombinase system is the FLP recombinase system of *S. cerevisiae* (O'Gorman *et al. Science 251*:1351-1355 (1991). If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein is required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al. Nature 385*:810-813 (1997) and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_0 phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyst and then transferred to pseudopregnant female foster animal. The offspring born of this female foster animal will be a clone of the animal from which the cell, e.g., the somatic cell, is isolated.

Transgenic animals containing recombinant cells that express the peptides described herein are useful to conduct the assays described herein in an *in vivo* context. Accordingly, the various physiological factors that are present *in vivo* and that could effect ligand binding, transporter protein activation, and signal transduction, may not be evident from *in vitro* cell-free or cell-based assays. Accordingly, it is useful to provide non-human transgenic animals to assay *in vivo* transporter protein function, including ligand interaction, the effect of specific mutant transporter proteins on transporter protein function and ligand interaction, and the effect of chimeric transporter proteins. It is also possible to assess the effect of null

mutations, that is mutations that substantially or completely eliminate one or more transporter protein functions.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the above-described modes for carrying out the invention which are obvious to those skilled in the field of molecular biology or related fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

<110> KODET, Stefan et al

<120> ISOLATED HUMAN TRANSPORTER PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
AND USES THEREOF

<130> CL000891

<160> 4

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 2782
<212> DNA
<213> Human

<400> 1
gtctcgtgta tggcgtggtt aaggttgcag cctctcacct ctgccttcct ccattttggg 60
ctggttacct ttgtgctctt cctgaatggt cttcgagcag aggctggtgg ctcaggggac 120
gtgccaagca cagggcagaa caatgagtcc tgttcagggt catcggactg caaggagggt 180
gtcatcctgc caatctggta cccggagaac ccttcccttg gggacaagat tgccagggtc 240
attgtctatt ttgtggccct gatatacatg ttccttgggg tgtccatcat tgctgaccgc 300
ttcatggcat ctattgaagt catcacctct caagagaggg aggtgacaat taagaaaccc 360
aatggagaaa ccagcacaac cactattcgg gtctggaatg aaactgtctc caacctgacc 420
cttatggccc tgggttcctc tgctcctgag atactcctct ctttaattga ggtgtgtggt 480
catgggttca ttgctggtga tctgggacct tctaccattg tagggagtgc agccttcaac 540
atgttcatca tcattggcat ctgtgtctac gtgatcccag acggagagac tcgcaagatc 600
aagcatctac gagtcttctt catcaccgct gcttggagta tctttgccta catctggctc 660
tatatgattc tggcagtctt ctcccctggt gtggtccagg tttgggaagg cctcctcact 720
ctcttcttct ttccagtgtg tgtccttctg gcctgggtgg cagataaacg actgctcttc 780
tacaaataca tgcacaaaaa gtaccgcaca gacaaacacc gaggaattat catagagaca 840
gagggtgacc accctaaggg cattgagatg gatgggaaaa tgatgaattc ccattttcta 900
gatgggaacc tggtgccccc ggaagggaag gaagtggatg agtcccgcag agagatgatc 960
cggatcctca aggatctgaa gcaaaaacac ccagagaagg acttagatca gctggtggag 1020
atggccaatt actatgctct ttcccaccaa cagaagagcc gcgccttcta ccgtatccaa 1080
gccactcgta tgatgactgg tgcaggcaat atcctgaaga aacatgcagc agaacaagcc 1140
aagaaggcct ccagcatgag cgaggtgcac accgatgagc ctgaggactt tatttccaag 1200
gtcttctttg acccatgttc ttaccagtgc ctggagaact gtggggctgt actcctgaca 1260
gtggtgagga aaggggggaga catgtcaaag accatgtatg tggactacaa aacagaggat 1320
ggttctgcca atgcaggggc tgactatgag ttcacagagg gcacggtggt tctgaagcca 1380
ggagagaccc agaaggagtt ctccgtgggc ataattgatg acgacatttt tgaggaggat 1440
gaacacttct ttgtaaggtt gagcaatgtc cgcatagagg aggagcagcc agaggagggg 1500
atgcctccag caatattcaa cagtcttccc ttgcctcggg ctgtcctagc ctccccttgt 1560
gtggccacag ttaccatctt ggatgatgac catgcaggca tcttcacttt tgaatgtgat 1620
actattcatg tcagtgagag tattggtgtt atggaggtca aggttctgcg gacatcaggt 1680
gcccggggta cagtcatcgt ccccctttagg acagtagaag ggacagccaa gggtggcggt 1740
gaggactttg aagacacata tggggagttg gaattcaaga atgatgaaac tgtgaaaacc 1800
ataagggtta aaatagtaga tgaggaggaa tacgaaaggc aagagaattt cttcattgcc 1860
cttggtgaac cgaaatggat ggaacgtgga atatcagatg tgacagacag gaagctgact 1920
atggaagaag aggaggccaa gaggatagca gagatgggaa agccagtatt gggtgaacac 1980
cccaaactgg aagtcatcat tgaagagtcc tatgagttca agactacggt ggacaaactg 2040
atcaagaaga caaacctggc cttggttgtg gggacccatt cctggaggga ccagttcatg 2100
gaggccatca ccgtcagtgc agcaggggat gaggatgagg atgaatccgg ggaggagagg 2160

ctgccctcct gctttgacta cgtcatgcac ttcctgactg tcttctggaa ggtgctgttt 2220
gcctgtgtgc cccccacaga gtactgccac ggctgggcct gcttcgccgt ctccatcctc 2280
atcattggca tgctcaccgc catcattggg gacctggcct cgcacttcgg ctgcaccatt 2340
ggtctcaaag attcggtcac agctgttgtt ttcgtggcat ttggcacctc tgtcccagat 2400
acgtttgcca gcaaagctgc tgccctccag gatgtatatg cagacgcctc cattggcaac 2460
gtgacgggca gcaacgccgt caatgtcttc ctgggcatcg gcctggcctg gtccgtggcc 2520
gccatctact gggctctgca gggacaggag ttccacgtgt cggccggcac actggccttc 2580
tccgtcaccc tcttcaccat ctttgcattt gtctgcatca gcgtgctctt gtaccgaagg 2640
cggccgcacc tgggagggga gcttggtggc ccccgtggct gcaagctcgc cacaacatgg 2700
ctctttgtga gcctgtggct cctctacata ctctttgcca cactagaggc ctattgctac 2760
atcaaggggt tctaagccac ac 2782

<210> 2
<211> 921
<212> PRT
<213> Human

<400> 2
Met Ala Trp Leu Arg Leu Gln Pro Leu Thr Ser Ala Phe Leu His Phe
 1   5   10   15
Gly Leu Val Thr Phe Val Leu Phe Leu Asn Gly Leu Arg Ala Glu Ala
            20   25   30
Gly Gly Ser Gly Asp Val Pro Ser Thr Gly Gln Asn Asn Glu Ser Cys
        35   40   45
Ser Gly Ser Ser Asp Cys Lys Glu Gly Val Ile Leu Pro Ile Trp Tyr
    50   55   60
Pro Glu Asn Pro Ser Leu Gly Asp Lys Ile Ala Arg Val Ile Val Tyr
65   70   75   80
Phe Val Ala Leu Ile Tyr Met Phe Leu Gly Val Ser Ile Ile Ala Asp
                85   90   95
Arg Phe Met Ala Ser Ile Glu Val Ile Thr Ser Gln Glu Arg Glu Val
            100   105   110
Thr Ile Lys Lys Pro Asn Gly Glu Thr Ser Thr Thr Thr Ile Arg Val
        115   120   125
Trp Asn Glu Thr Val Ser Asn Leu Thr Leu Met Ala Leu Gly Ser Ser
    130   135   140
Ala Pro Glu Ile Leu Leu Ser Leu Ile Glu Val Cys Gly His Gly Phe
145   150   155   160
Ile Ala Gly Asp Leu Gly Pro Ser Thr Ile Val Gly Ser Ala Ala Phe
                165   170   175
Asn Met Phe Ile Ile Ile Gly Ile Cys Val Tyr Val Ile Pro Asp Gly
            180   185   190
Glu Thr Arg Lys Ile Lys His Leu Arg Val Phe Phe Ile Thr Ala Ala
        195   200   205
Trp Ser Ile Phe Ala Tyr Ile Trp Leu Tyr Met Ile Leu Ala Val Phe
    210   215   220
Ser Pro Gly Val Val Gln Val Trp Glu Gly Leu Leu Thr Leu Phe Phe
225   230   235   240
Phe Pro Val Cys Val Leu Leu Ala Trp Val Ala Asp Lys Arg Leu Leu
                245   250   255
Phe Tyr Lys Tyr Met His Lys Lys Tyr Arg Thr Asp Lys His Arg Gly
            260   265   270
Ile Ile Ile Glu Thr Glu Gly Asp His Pro Lys Gly Ile Glu Met Asp
        275   280   285
Gly Lys Met Met Asn Ser His Phe Leu Asp Gly Asn Leu Val Pro Leu
    290   295   300
Glu Gly Lys Glu Val Asp Glu Ser Arg Arg Glu Met Ile Arg Ile Leu

305 310 315 320
Lys Asp Leu Lys Gln Lys His Pro Glu Lys Asp Leu Asp Gln Leu Val
325 330 335
Glu Met Ala Asn Tyr Tyr Ala Leu Ser His Gln Gln Lys Ser Arg Ala
340 345 350
Phe Tyr Arg Ile Gln Ala Thr Arg Met Met Thr Gly Ala Gly Asn Ile
355 360 365
Leu Lys Lys His Ala Ala Glu Gln Ala Lys Lys Ala Ser Ser Met Ser
370 375 380
Glu Val His Thr Asp Glu Pro Glu Asp Phe Ile Ser Lys Val Phe Phe
385 390 395 400
Asp Pro Cys Ser Tyr Gln Cys Leu Glu Asn Cys Gly Ala Val Leu Leu
405 410 415
Thr Val Val Arg Lys Gly Gly Asp Met Ser Lys Thr Met Tyr Val Asp
420 425 430
Tyr Lys Thr Glu Asp Gly Ser Ala Asn Ala Gly Ala Asp Tyr Glu Phe
435 440 445
Thr Glu Gly Thr Val Val Leu Lys Pro Gly Glu Thr Gln Lys Glu Phe
450 455 460
Ser Val Gly Ile Ile Asp Asp Asp Ile Phe Glu Glu Asp Glu His Phe
465 470 475 480
Phe Val Arg Leu Ser Asn Val Arg Ile Glu Glu Glu Gln Pro Glu Glu
485 490 495
Gly Met Pro Pro Ala Ile Phe Asn Ser Leu Pro Leu Pro Arg Ala Val
500 505 510
Leu Ala Ser Pro Cys Val Ala Thr Val Thr Ile Leu Asp Asp Asp His
515 520 525
Ala Gly Ile Phe Thr Phe Glu Cys Asp Thr Ile His Val Ser Glu Ser
530 535 540
Ile Gly Val Met Glu Val Lys Val Leu Arg Thr Ser Gly Ala Arg Gly
545 550 555 560
Thr Val Ile Val Pro Phe Arg Thr Val Glu Gly Thr Ala Lys Gly Gly
565 570 575
Gly Glu Asp Phe Glu Asp Thr Tyr Gly Glu Leu Glu Phe Lys Asn Asp
580 585 590
Glu Thr Val Lys Thr Ile Arg Val Lys Ile Val Asp Glu Glu Glu Tyr
595 600 605
Glu Arg Gln Glu Asn Phe Phe Ile Ala Leu Gly Glu Pro Lys Trp Met
610 615 620
Glu Arg Gly Ile Ser Asp Val Thr Asp Arg Lys Leu Thr Met Glu Glu
625 630 635 640
Glu Glu Ala Lys Arg Ile Ala Glu Met Gly Lys Pro Val Leu Gly Glu
645 650 655
His Pro Lys Leu Glu Val Ile Ile Glu Glu Ser Tyr Glu Phe Lys Thr
660 665 670
Thr Val Asp Lys Leu Ile Lys Lys Thr Asn Leu Ala Leu Val Val Gly
675 680 685
Thr His Ser Trp Arg Asp Gln Phe Met Glu Ala Ile Thr Val Ser Ala
690 695 700
Ala Gly Asp Glu Asp Glu Asp Glu Ser Gly Glu Glu Arg Leu Pro Ser
705 710 715 720
Cys Phe Asp Tyr Val Met His Phe Leu Thr Val Phe Trp Lys Val Leu
725 730 735
Phe Ala Cys Val Pro Pro Thr Glu Tyr Cys His Gly Trp Ala Cys Phe
740 745 750
Ala Val Ser Ile Leu Ile Ile Gly Met Leu Thr Ala Ile Ile Gly Asp
755 760 765

```
Leu Ala Ser His Phe Gly Cys Thr Ile Gly Leu Lys Asp Ser Val Thr
    770                 775                 780
Ala Val Val Phe Val Ala Phe Gly Thr Ser Val Pro Asp Thr Phe Ala
785                 790                 795                 800
Ser Lys Ala Ala Ala Leu Gln Asp Val Tyr Ala Asp Ala Ser Ile Gly
                805                 810                 815
Asn Val Thr Gly Ser Asn Ala Val Asn Val Phe Leu Gly Ile Gly Leu
            820                 825                 830
Ala Trp Ser Val Ala Ala Ile Tyr Trp Ala Leu Gln Gly Gln Glu Phe
        835                 840                 845
His Val Ser Ala Gly Thr Leu Ala Phe Ser Val Thr Leu Phe Thr Ile
    850                 855                 860
Phe Ala Phe Val Cys Ile Ser Val Leu Leu Tyr Arg Arg Arg Pro His
865                 870                 875                 880
Leu Gly Gly Glu Leu Gly Gly Pro Arg Gly Cys Lys Leu Ala Thr Thr
                885                 890                 895
Trp Leu Phe Val Ser Leu Trp Leu Leu Tyr Ile Leu Phe Ala Thr Leu
            900                 905                 910
Glu Ala Tyr Cys Tyr Ile Lys Gly Phe
        915                 920
```

<210> 3
<211> 126512
<212> DNA
<213> Human

<220>
<221> misc_feature
<222> (1)...(126512)
<223> n = A,T,C or G

<400> 3

```
ttggatgaga tctaaagcat tattaagagt ggggagtgca aagaagaaac cctcatttca 60
aagatgaatg agaataatgg catgtacaaa ggtcctgggg tggacagtca cttggtataa 120
tccaagagtg aacctgaagg ctattgttgt tgaaatgtaa taagggagag agtgacggga 180
tgaaggggga tgagtgggaa gcagtgaatt cctgcaaggc tttgaaggtc atgggaaaga 240
atttggtctt tatatcaaga gcaagagaag actactaaag ggcttcaaac aggggagcga 300
tatgcttaag tctgtttgtt tgtttttta aaaaaagatt acggtggcta tatgaggaaa 360
gtggaattga gaactagcga gagttggagt ggtgagctcc attaggaggc tactgaagta 420
gattcatgag gtaaggagtg atggtggcct gggctgggat gatggtggta gaaatggaga 480
aagagttgat aggatttagt gattggataa gggacagaag agagatgaag gctttcagac 540
taacatctgc tttctaacat gagtaactgg gtggctgaag atgctatttt ctgagctggg 600
aaacaggaga aaaaggagca aatatggggg atgaagactt tgagtcttta aggtgctgta 660
caaacacaaa tcagcattcc tttattacta agggtatccc acacagttgt agcagaggga 720
gaaagatcgc ccccccccca ctttttttt ttttttagct attccatggt attttcattc 780
tcatcccacc caaatgaggc agtgagtggt aagatgagta tataatagtt tcaattgcat 840
ttcatcccat tcttctgagc tcaagctcac cttttagtgg tttgaggcca gtagatgaag 900
ctgcatatca cccccaaaat cttgtctcta gtttaacaaa acttatttga gagacatttg 960
catgttttat taataatgat ttttaccact tgttcctttc catgtttggg tttgaaattt 1020
gagtggctgg cggatgatca tcttcctgtt actgcctgct taaactgctc ataagcaggt 1080
tttactggag ggctcagagc tgctgtgaac ttggtcttgg gcacaactta catggcctct 1140
gtttggctat ggggtgggtg gcattcacca tttatcaact cttttgattt cccaagctat 1200
ctcagaatta tagcttgcct ccagaagtct tgcattcggg gaggaagttt ctttccaagg 1260
gagctcagtt ttcaaggttt attgctctgt ttaatggatg agatctaaag cattattaag 1320
agtggggagt gcaaagaaga aacactcatt tcaaaatcga ttgagaataa tggcatgtac 1380
aaaggtcctg gggtggacag tcacttggta taatcctgga gtgaacatga aggccaagga 1440
```

aatatgtata cattaaacag agcaaggttt tcaattttct ggggactagt ccatgaaaat 1500
tcaattcaat atactctctt gcaaacctat gttatccaag atactcaagt ataatgacaa 1560
cagggtaagg aagtccgaac accccagaaa cagtataaat gggcatgaag attcaggtta 1620
tacatggcct attttaagtt gcttcttgag aactctcaca ggtaatacca gtttgggaga 1680
caggacttga aggctattgc tgcatttcca tccccagtat tcccagctat ttcaagccat 1740
ttttcaacgg agtctccacc agatggtttg gaggacagag cagctatttg tgcctcccat 1800
tgacatctat ttttccaagt gagagactgc cccatatgtt agtgcaatat gtcactggag 1860
gtgaagcatc agttgtattg gtgggaacct gccgtttgct gtccccttt tcctcatgcc 1920
ttttcctgcc tctctgatct tttctaggtc tctggcctat caggaggaca actggtgctg 1980
caatagaagc cagtggctaa gtctcgtgta tggcgtggtt aaggttgcag cctctcacct 2040
ctgccttcct ccattttggg ctggttacct ttgtgctctt cctgaatggt cttcgagcag 2100
aggctggtgg ctcaggggac gtgccaagca cagggcagaa caatgagtcc tgttcagggt 2160
catcggactg caaggagggt gtcatcctgc caatctggta cccggagaac ccttcccttg 2220
gggacaagat tgccagggtc attgtctatt ttgtggccct gatatacatg ttccttgggg 2280
tgtccatcat tgctgaccgc ttcatggcat ctattgaagt catcacctct caagagaggg 2340
aggtgacaat taagaaaccc aatggagaaa ccagcacaac aactattcgg gtctggaatg 2400
aaactgtctc caacctgacc cttatggccc tgggttcctc tgctcctgag atactcctct 2460
ctttaattga ggtgtgtggt catgggttca ttgctggtga tctgggacct tctaccattg 2520
tagggagtgc agccttcaac atgttcatca tcattggcat ctgtgtctac gtgatcccag 2580
acggagagac tcgcaagatc aagcatctac gagtcttctt catcaccgct gcttggagta 2640
tctttgccta catctggctc tatatgattc tggcagtctt ctcccctggt gtggtccagg 2700
tttgggaagg cctcctcact ctcttcttct ttccagtgtg tgtccttctg gcctgggtgg 2760
cagataaacg actgctcttc tacaaataca tgcacaaaaa gtaccgcaca gacaaacacc 2820
gaggaattat catagagaca gagggtgacc accctaaggg cattgagatg gatgggaaaa 2880
tgatgaattc ccattttcta gatgggaacc tggtgccccc ggaagggaag gaagtggatg 2940
agtcccgcag agagatgatc cggattctca aggatctgaa gcaaaaacac ccagagaagg 3000
acttagatca gctggtggag atggccaatt actatgctct ttcccaccaa cagaagagcc 3060
gcgccttcta ccgtatccaa gccactcgta tgatgactgg tgcaggcaat atcctgaaga 3120
aacatgcagc agaacaagcc aagaaggcct ccagcatgag cgaggtgcac accgatgagc 3180
ctgaggactt tatttccaag gtcttctttg acccatgttc ttaccagtgc ctggagaact 3240
gtggggctgt actcctgaca gtggtgagga aaggggaga catgtcaaag accatgtatg 3300
tggactacaa aacagaggat ggttctgcca atgcaggggc tgactatgag ttcacagagg 3360
gcacggtggt tctgaagcca ggagagaccc agaaggagtt ctccgtgggc ataattgatg 3420
acgacatttt tgaggaggat gaacacttct ttgtaaggtt gagcaatgtc cgcatagagg 3480
aggagcagcc agaggagggg atgcctccag caatattcaa cagtcttccc ttgcctcggg 3540
ctgtcctagc ctccccttgt gtggccacag ttaccatctt ggatgatgac catgcaggca 3600
tcttcacttt tgaatgtgat actattcatg tcagtgagag tattggtgtt atggaggtca 3660
aggttctgcg gacatcaggt gcccggggta cagtcatcgt ccccctttagg acagtagaag 3720
ggacagccaa gggtggcggt gaggactttg aagacacata tggggagttg gaattcaaga 3780
atgatgaaac tgtgtaagta accttcctgt attctgcccc tccctgaccc catcttttgc 3840
catctctttc tgtctttctg tactgcactt tacaacattt ccttgtgttt gtgttaatgt 3900
caaactttgg ttccatcaca ggtatgcagg atcagcagac accactggac aggttctgct 3960
tccaaactct tcttcagttt tctcacttta aattgtttct gggcaaggaa tcctgtgaca 4020
agagctaagg acacaaaaca ttttcttctc tgaaacacaa aatgatagct ggtggagctg 4080
tgggatgaca gaagttttgt gatatcagat tttggagaat tcttgtgact aagaaggact 4140
agagaactgc ttgggcctct tcttcctccc ttcctcatat gaagggtatc tatgagcttt 4200
gaaaccaatc ctttccattc tgggcagcaa tagcccatca gaacattcta aagaaaacaa 4260
gtggcattgg ctttgttccc tggtactata ttgccagtct cactgtgtaa ccagattcca 4320
ggcacgtctt ctttaatttg gaaattgcaa aattgataga aatttagcaa tctttttaaa 4380
tgaccataga ctatttaatg gtgtgaggct tgcccagcct agttgaattg agtcagtatg 4440
gtttggatac tggaaagtat cttggagaag cagagctccc agggcagtgg ctacttgtct 4500
ttagtcacag gtctaagctc caaaatctgg tgaagcagtg aaggagaaac atcctaggaa 4560
ttgtgggagg aaatatatct tctgtgtggt cctctctttt cacagtctag gactctcctg 4620
aagtacctct tcttgggcta ctgccccatt cagcccttca gaaactgtgg gtattacact 4680
tctgtcacct ctattaccct aaggcctctg cccattgaac cctcttgcaa attggttatt 4740
ctgtcctttt tccagttgga tagctttaaa agggaaagca gaatgacttt cctcaggatt 4800
tgtagcttat gagaaagtag actttcttgg gtggcctaga aggttggaga agacaaacgg 4860

gaacttcctc tgaatgactg aacatatcca caaataataa gcgtggcagg agatggtgtg 4920
aagagtaaaa ggagcatata ggaagttgtg tgtgtggggt gtctgtttca agaacctgct 4980
aattatacct tcagtaagaa atgaagccat acaacctcta gaagaggagg aggaaggaac 5040
tcatggaaaa gtggggagcc atagaagcta gggagaggtg tcctaggagt gcttctgccc 5100
aggtccagcc atgagacaga gctcaaaaag agctgggcac tgctggtgac agaactgagt 5160
gacccggggg atcctgcatc tgttcttact caatcccttc ttaataatgt gacttggggc 5220
aggtcattta ttggttctgg aacttaactt tctgatatgc aaactgggaa taacaatact 5280
ttccttgcct ggaggcaagg tcagtccttt ttgcagttcc ttccagctct aagattttct 5340
gaaccataga cataagcact cagtgtaggt catattcgca cttgccaaaa atggatcagg 5400
gaatattgtc tcctgaaggg aaatggccat tgacaaattg atttattaga gctctgttta 5460
gtcattttgc tgggaaggat aatcatttgt taacgtaagt agaaacctgt gccttctgga 5520
gaatactatc catttatatg tactctgggg agagtgttta tacatacaaa tgaaggacag 5580
ggcttcactg ggaaaacaaa ctccatggaa tttcacatga ttatcgcgat gtcagtgtgg 5640
aagaagatat ggtaaggcat taaatgacat taagaccaca aaatttgcca taatttgacg 5700
gacttgtggt tcttctgatt cagaaccctt tctacccatg tcacggatag gtagtttttc 5760
agagatcaga ggcttagttc attctattaa tttcctcatt ctattaataa tcaattatgc 5820
acctagggtc tctgaatacg actaaacctt cctcaaactt atttgcattt tcagtttgta 5880
taatatcttg gtgcaaatga gcctcgcaaa tgatcacttc tgggtaatac tcattctaaa 5940
ggtatgtcaa ccttgagaat tctggtctag atattctagg gtttggtgaa caaatctatg 6000
ttcccatcca tccctttca tttatttttt agacttcatt cattgcagaa taatgagtcc 6060
aaaacctgct catctgttct cacgtggcac ccctattctt gatattttaa attgcaattt 6120
tacaactaga ggcagtatta cggagcagaa aaatcgtggg ttctaagtac tctgggttag 6180
gattctggct ccactactga tttaataatg tagtttgggg aaattttatt aacctatgaa 6240
attatttcct cattggcaaa atggggataa taatatctct cttgcagggc cattatgacg 6300
attcaaggta ttgtatgcgg tgtacctggt acacggtata tgctcaggaa acaagactct 6360
tcatagtaat attgacgaat taacaatatt cttcagaaga cactgtggag ttgtttaggt 6420
tacttggctc tttgtgtgac cctaagtaat gagcatgcca gtttggggtt actatgaaga 6480
gtacttacct aaactcataa aatattagag ctagaaagga ccttagaata tcttctgcag 6540
tcatggttct taaattttaa tgtgttgctc aatcatccag ggatctcact gaagggcaga 6600
ttaggatcca ggaggtctag gggagggatt gagattccgc atttctaaca agttctggat 6660
gctgcgggcc ccaacttaga ggtgaaaggt tctgaagctc ttgaccaaac caggagaccc 6720
agcaaagaag tggtttttca gacaacttgc ttaattgaat aatgattgtt tgctctttaa 6780
ttccaacttt caatgccaat ttagcaagaa ccagaggctg tgctaattgc cacaccagtc 6840
tggaaaccga aatggatagc ttcagggtac ttggacaaag ttggaacatc tgctttctaa 6900
tctctccctc tttgtatagc tttatttgcc taccaagcct ggtagtattg aaaatctgcc 6960
ctcactatac tcccctaaat ataatcaagt tgaggccagg cctgtgctct atcaataata 7020
taggatccac gaattcacat gtttggtttt atgctttact tcttcaaagg tgcttttagc 7080
agcatggaag aatggaaaag cacgagcttt ggaatatgaa agcagatgtg aatccatcac 7140
ttaccagtaa cttttaacaa gtcacatcac ttttctgagt accaggtttt tgttggacaa 7200
cagaaataat attctctatc cttcaaggga atactaaata taagtatgag aaaaatgcac 7260
agtgccttct cgtagatggt gttcagtcat tcaacaaaca tttgttagat atttgctatg 7320
tactagctac attactaggc actggggtta aataagtgaa taagacaagc tgacatttca 7380
gcgctcaagg atcttactgt caagtggaga ggatcaaagg gtacagacaa atcaaggaac 7440
gtgagagaag tggtatggct gagatggatt gaataaagga gcaatgagag ctccctgcaa 7500
tgtgtgtggt accactgagg attctaaatt aaccttcatt aaggacttag tagtgacaga 7560
ggtgaagtgg ggataggtac atgattaatt tacatccata ttacaatgaa accttaacat 7620
ttaagaggga tattattgat gtcttcatga tccagaagaa tcctcacctt tgcaaccatc 7680
actatagtca cttcttgaga attatggcct ttaagactgt agcatgcaat gacaaaacct 7740
cacagaggta tgggttctgc ccgcacacta atttcactca ttaaacaagt gactggctcc 7800
tatatcccag gctctcagca cgcctttgca aaataacaga ttattgcagc tcttggacct 7860
ttgatgcctc tgggaatagt caaagccaca gatgtcaaat atgtaaatgc caagatctat 7920
tataattaaa tagtgcaggc ctccttcaaa gaaaaaaagc atgttggctg tgctgcacgt 7980
tctccaacca aatcagaatg ttaaagctcg aaggtatctg acctcccatt ttttaaatta 8040
tgaagatgaa attcagaaag ggaaggtaac ttatccaaga ttacatggct agctatgata 8100
gaaagttaga gttggaaagg acgttagaaa gtgagggttt gaaaggactt tagaagctgc 8160
ttattcaatg ttctctctgc cctttcccat cttaggcttc tccattttac ttttatccat 8220
caataaaatg ttaacttcaa aaagaatatg gcaattcttg ggtaaaagat gctctggaag 8280

tgtgagtccg ggagtattat gtgactaatg tcttaactaa gaataataat atattatgga 8340
ctagttttaa tctcttgttt caccttgaac tgttcaggaa ggaaaatagc ccacggaaat 8400
tttttaaaaa gtctttctct atctgaattg agaaaaggtg acaggcatag ttggaacatc 8460
ttttaggcag tgctggtgaa cttcaggcta ggccttgttc catgaaataa taaaaatttt 8520
caaaataatg cagaccattc ccttccaggg atgctttctc tgtaatgttt taaccccaag 8580
aaatctttct gtaaaaatct ataaaaatct ggagtgttcc aggatacaat ttgcacattc 8640
tccaatttaa ctaaaacaca atcgattttt tgttttcttt ttctttggct tagcaaggtt 8700
ttaagatagt ctctttctgg ccacagaggg agatgatttg cctctagaat acccttttctg 8760
tgcttgagag agtcacaaga ctgcaagctc atggaggatg agagtcaagt agaggtggtg 8820
acatctctcc cttggccaac atccctctct ttctctttcc ttctgccttc agtggcagta 8880
gcaaaagtcc tccttctctt taggtagaca gtcagccact acaactgtgg cttcctgaaa 8940
tcctcagtgg agctatgtac ttggcacaga tttgtcttga agaagggact ccatttctga 9000
gccagttgtt gaatggggat acttagcagt acagtgaggc atttccagta ggattgttca 9060
accacaattg cccactttcc aggcccaaag gaataattga aggctatgta gacttttttt 9120
ttttttttt ttttttttt tttgagatgg agtctcgctc tgtcgcccag gctggagtgc 9180
agtggcacat ctcggctcac tgcaagctct gcctcccggg ttcacgccat tctcctgcct 9240
cagcctcccg agtagctagg cctaatatat atatattata catatatatt tatatttata 9300
tatatatata ccaccacgtc cggctaatat atatttatac ttttttttt tagtaggaaa 9360
ggggtttcac catgttagcc agtatggtct cgatctcctg acctcgtgat ccaccagcct 9420
cagcctccca aagtgctggg attacaggcg tgagccaccg tgcccgacca tgctatgtaa 9480
acttttagc agaagcttta gctattgtgt cccgaagggc cccaggtcat gatgaaatgt 9540
ctttttttt ttttgtctct tttcttctta attactgaga ctgtcaaaga atatgtcaaa 9600
gcatgacata ttccaactcc aggatccata aaacacccca agttctgtgg agaccctatc 9660
acatctgcaa aactctccag gaagtccaga gccctcctgg ttaatttgtt ttagggacta 9720
ggcatgcggt atcccctgac aacactggat cagcaattct cctacctaag tcagtcccac 9780
accatgtgca gcagagtatc cagtgcccct gccctggtct gctcacattg gtttgctctc 9840
cagaataata attcctcaat atccacaaga gattgattcc agaactactc cgaggatacc 9900
aaaaatcctc agatgctcaa gtacctggta taaaatggca cagtatttgg catatgacct 9960
aggcatattc tctcccatat actttattta tttatttatt tcgggacaga atctcattct 10020
gtcgcccagg ctgtcactcg cttattgcaa cctctgcctc ccaggttcaa gcaattctcc 10080
tgcctcagcc tcctaagtag ctgggactac agacgcatgt caccacgcct ggctactttt 10140
tgtattttta gtagagacag agtttcacca tgttggccag gctggtctca aacacctgac 10200
ctcaagtgat ccgcccacct tggcctccca aaaagctggg attacaggcg tgagctacca 10260
cgtccagccc cccatatact ttaaatcatc tctagattac ttataatacc taatacaatg 10320
taaatgttat atagttgttt taatgtattg ctttttttat ttgtattgtt ttttattgct 10380
gtattatcct tttttatgtt ttatttttttc aaatattttc tacccgtggc acccacagtt 10440
ggttggtgga acctgcggtt ggtggagccc atggatgtga agggctgata gtatgagaaa 10500
actcagaggt gcagagttgg agagcacatc ggggagaatg tcagcatggg ttaaaaaaga 10560
cacactgtgg ttggagatga tcacatgaat ggccacttca aaaatgaatg ggtctcatcc 10620
tcaaagcagg ctctcctggg cactgcttgg gaaggtgcta attggagctt caggcaacaa 10680
taataagggg atacaggtgg ggatcctgcc atgggcgtag cttactttct ctggactctt 10740
ctgggtctta aggccagttt cctcatccac tcaaaagaat gacagcaagg tgagcaaagc 10800
aaggcaggta aatgaggagg actctttctg gctgtccaac ttttcatcaa cttcccaaag 10860
gtttttggat gggacatgag cactcattcc ttctccaccc tttagctagg ccctgtcaac 10920
tccaggagga aggtagaaga ggtcagagct gtggtctttc acttattcaa gatgtttcct 10980
tagtgttttg tgtttgggtt ttttttgttt ttttttttt gacagagtct tgctctgttg 11040
cccaggctgg agtgaagtga agtggcataa tctgagctca ctgcaacctc tgctttcgag 11100
ttcaagcgat tctcatgcct cagcctcctg catagctggg actacaggca tatgctacca 11160
tgcctggcta atttttgtat ttttagtaga gacggggttt tgccatgttg gccaggctga 11220
tctcaaactc ctgacttcag gtgatccagc caccttggcc tcccaaagtg ctgggattac 11280
aggtatgaac cactgcacct ggccccttat tgttggtttt taaaagagaa actaagctgt 11340
gcttccagaa cccagtttga gaaagtttga agacctggca tagagccagt gacatataat 11400
tgttagttga agaaagagag ctccttgatc tgcaaataga gcacggcccc atatttaaat 11460
tctgcacatt ctagaagcat tttgcaagaa tcaaatgctt tgaggatttt gctaaataac 11520
catggaggaa agcactagac aaatattttc agatggcatg agagttatca ttcataggaa 11580
ttatatttcc actcctacca cttactgggg acccaagtaa gaaattactt ggataagcag 11640
aggagaattt aaagttgaat gtggtggaac ttattatgga aaaaatatgt ttttctgaaa 11700

actggatatg tgtatatata taagttcagt tgtcattttg gaaccatcct tactcttcct 11760
agctaaggat tagcatacat aggtgcaact tgactaactc tgcctggacc caattcagtt 11820
acctttggt gggtagggtt catgaagaag cagttatttg tggagtgtat agaaaccact 11880
ctattgtagg ttctttagtt ggtactttca aaataagtga catccaaata gtaacttaat 11940
attccaaata tggctgcaaa acaaattgtc gattatggat gactactact gccatctctc 12000
cataccagtc catcttctgc caggctgttt ggtcttgatt tgtcgacctt ttaggtttct 12060
ccccatgtat tccacatgac cttcaccaac cccacttcta tctccaaacg tctttctgag 12120
ttgtggggat gcagatgtat tctgccacca tcacaagggc taaccgagcc ctggctgcgg 12180
atcttcattg ttgttcacat tatttccatt cttacaccct acttcatgtt tgtacactat 12240
tttcttacat ttgctgtctc ttctaaacat tctttgctgc atccactttt tctctatttg 12300
tgctctaggt gctgcagagg ctaatgctgg gtttcctttc attcctcctt gcactcagca 12360
cctcccttct caattccttt tgccatgtct ccactttaaa tcttaaccta ctccagatag 12420
tcttttcctt cacactattg gcatctgtgc ttgggttgct ttcagtctat tctctgatct 12480
atgatttctt tgcatgatca agaaggtgcc atgaaaggat cccttaagaa agcctgtcat 12540
ttagccagaa cgaactagct tcatgatagc accaggaaga ctgatatctc ccaggaaaca 12600
aaccactcat ggtggtgctc tttttgcctt cactatgaag tgtttgtctg cctgtatgtg 12660
aaaacgagag ggtttaattg taaggatgca gcacagattg ggactggcat cagaaagcca 12720
ttggggactg aggtagctct agagaccgct ttctgtctcc agtgctctcc ctcctgggtg 12780
acatgttttc tgtctcctgg catctctgct tctctctatg ggcttcttta ttatttgcag 12840
cttgcaatgg taccccaaag tcctagctca tggctcctct ctgcatatat gctttctgtt 12900
cctacccaca aagctctttc tattcttcta gtttaaattt tcaagagaag aaatctgatt 12960
ttttttaac ctggtcatgt caaagaccac tgaccacata tgagctggtt gccctgtgtc 13020
aagtgccccc ttctcccacc ctcttcccct ccccatctgg tctgtcataa ctgaatgatg 13080
gagtgggaaa ttgaaattgc catgggaatt ccatgataag ctatctaaac agttttatct 13140
ataagtggta gacagagtca cttagaaggg agtcccaggt gagacaggca cctgtcaact 13200
ccaaactggc acacattcta aggtctgcaa caccccagag agagcactga ttttgtagtg 13260
gcctgtactg gggcggtagg ctggagaatg ggagaaatag ccacttcaga atcccccagc 13320
ccaaatgcat caagctcact atagactctg cagccacgat tcagctggct tctgctcaga 13380
tcaacagaaa acattcttag tgaatgatgc ttgtggcaca tatctcaagg ctaccagggt 13440
catttcttcc catttacttt ttctctgatc tatcctctcc aggacactag cgtcagaaga 13500
taatcttccg tcgttttcag gtacactatt tgggtactga gtcactttca aagcctcttt 13560
ctgggtttgg atttccagag cagcctgtgc tgtaaagcaa gacagaaagc ttccctgcca 13620
ttcatgcctg ccagggatag aatgacagta ctcctgaggc tctccctccc caccctccc 13680
ctgctggaca gctgatctgc tggactcagc cagagccagc aggcaccccc tctttatcct 13740
aggagctgca aacttgatgc ctttccagga aatccccaga agctggagta tcctcatcta 13800
catgtggcac agtgtatggt tgtgtcaggt gctcatgtcc cattgcatag gactggggtg 13860
gaaaataggg accgtccttt tgtgtcagct ccagtcaatg agtagtggcc atccaggggg 13920
ccatcttgga aaggacttgt gaggctgtat ctgcgctcag ttgtagatgt gagaagaaaa 13980
ggccaaatat ctgccaatcc tagtcctggg attcaagata gaaagaactg catggagtga 14040
agaaactagg agtctccatt tcactgagat gcataagaat gaaattattg tcactatttc 14100
ttcaatactg ggccaatcct aataagaaaa ccctttttga gtctctcttt tctttatcct 14160
acatataaca cagaagcttt ttctattccc tggatgaacc cacagggaca gaaattcttg 14220
ttggacaggt gaagcagata atttctttat cagactagaa tcttccagaa gcactgctaa 14280
cctagtgagt tttgtactct agacaggtgg ttctcaagcc agctccccac cgcaggcctt 14340
tttcatggtc tgcccctccc tgtggaaccc atgttttagg ttattagctg ataattggat 14400
ttctattttt tctcataaaa tacagcaaaa gatagctagt gatattatga tgagttaatg 14460
taattatagc caaagcagag agaaacaaca ttttaattaa cctgtgtgga ctgctggaag 14520
aatataaact ttctattttg ggggttgagt agagacagaa atgaacacag ccaagggctg 14580
actgtcagag gacatttaac tgatgtaaaa tgctttgaaa ttattgggca ctcattgttt 14640
aaagttgttt ttgatgatgg taactccgta aggggatcag aacatgctgg aaagaatggg 14700
cacagctttg gttacctggg ccttaccact gttattcagg cctctgagaa agcttactat 14760
tgttgttatg tttcttacat aataaaactt ctaatatttg tatgaaaaca tagaattcca 14820
cttttaaaga tgtaaggatt ttgtcatacc attagggtta ctatgatcac ttgattctag 14880
gtctaagaaa tattaagtaa tttacccgcg aacacagagt tttaagggta agtatcaaaa 14940
ccttgatctt ctaataccac atattctcac tcatatgtgg gagctaaaaa tattgagctc 15000
aaaaaggtag agagtagaat tgtagttatt agaggatgcg aaggaggata gggagaggtt 15060
ggttaatgga tacaatgtga agttatgtaa gaggagtaag ttctagtgtt ttgtagcact 15120

gtagggtgaa tatggttaac agtaatttag tgtatattta aaaaaaaata gacaggattc 15180
tgaatattca caaagaaatg ataaatattc agctgggcgt ggttgctcac gcctatattc 15240
ccagcatttt gggaggccga ggtggatgga tcacctaagg tcaggagttt gagatcagcc 15300
tggacaacat ggtgaaaccc cgtctctact ataaatacaa aaaattagct gggcatggtg 15360
gcgcacacct gtagtcctag ctacttaaga ggctgaggca ggagaatcgc ttgaacctgg 15420
gaggcagagg ttgcagtgag ccgagatcac gccactgcac tccagcctgg gtgacagagt 15480
gacactctgt ctcaaaaaaa aaaaaaaaaa gaatgataaa tatttaaggt gatagatatg 15540
ctaattaccc tgatttgatc attacacttt gtatacatgt gtcaaaatat cactctgtat 15600
ccatacatat gtataattat tatgtgtcaa ctaaaaataa aaggaaaaaa atcatttcag 15660
tgtatttaca aaacatatgt aaccattaag aataatgttt taaattatat ctaagggtgt 15720
gataaaatta cagtataaga ttgtgcttga aaaagtgcaa taagaagtaa atatgtacag 15780
atgagaaaaa gtgcaaagaa ctaagtccta agcagactat acctttccta ctgcatggta 15840
cttctctggc cttttgcttt gaaagatttt gcacccagca tggcaagtgg ttagcagagg 15900
cagccattct cacttgtgcg ttggctttgg gagccatata tgttgttcag ctgggtgtgg 15960
agtggaaagg ctgcatgttg tattaatgca ttgttaagaa cctctaagag tgatttcttt 16020
tgggaagtga gactgacggt ccgaatggtg gaaagacaac ttttaatctt ttactttaca 16080
ctttgtgcac ttttaaatgt ttaacatgag catgcatttc tttaataata aaaatacaaa 16140
aaaattttag ccctagatct tctgatttta aactgcatat tctttctatt gtgttacata 16200
ttttagcatg agaataaggt tatgaagctg gaagtagcag gctccctttt cctcatatgt 16260
aggaagttaa gaatgcattc tacgtttctt ctttaaggag ttggcttctt tccttttaac 16320
ataggggtaa ctgggcccag ggagtttggc aagggccaaa taaagtcctt aatgcccagc 16380
tcagaaatct ggattcacca tccttgactg ctggctccaa cccaccctca cctgagctgg 16440
tctgcagagg attcttgttt gtgtcacttc atcaccagca actaccgaca gatgatgctt 16500
tggcctgctg cctgggtaac agggcgaggc tggctcagga ccatgttttc agatcagggg 16560
acctcctttg atgccatgtc catggtgtcc gagggcagcc aggatcaagg gctagacggg 16620
gcagtgatga gatgagagca ggaggggctc agctgcagcc ccaggagagc ctatgccagc 16680
cctgttgacc aaggaggaca gaagcaacag gagagcggag gcagaggggt gagtgtctat 16740
cgctcaatgt ataatcggca gacatttggg gagctcatac tgtgggctaa gcacagggaa 16800
gaaaggcaca gtccctgtcc tcagggaggt cacagttgat agggaagaca agcatatgtg 16860
ctagctgcta tagaaggggg aaccactgag ggctgtggcc acacagaggc aacacccccc 16920
tcttgttttt ttgtcaggga ttcagtttgg cgtcattaga agtgacttgc acaacccccc 16980
cctccagtca attcagaagg acttgttaag caggaatgat gaattagctt cagcttgtgg 17040
ggcacacaca gatggaagta taaggtggcc tcaggagtaa gtaaatcccc atgcaagctg 17100
tgtccttaga ccagagcagc acccggttct tccccatttc tagtaaaggt gcctcacaca 17160
ccaccaggac acaatttatg cctgcagaat gaatgaatga atgaatgagt gaattcctgg 17220
aacctcttct gcttatgtgc cacaccaggt tgcagcaagc ccagggacac ctgggactgg 17280
aattgggctc tcaggtgtaa ggaccaggga gcacccacca ttttgcattc ttcagccctt 17340
cctcctctcc tgtcccagct tcagcaatat ccacagagcc ctctgagcaa ctctgagcct 17400
ctccacagcc tgacgcctgc ctgggcacca gctcttcaga gggtgtttct gtgctgctca 17460
gctacctctg agcctgggct gcctttgatg ctcaggagac accctgtaat tcaattaagc 17520
cttctctcca gggagcatgt aattatgtcc tatctgggcc ttgtaatgac agccccctgc 17580
cactctacag ggagttgccc tgctcagctg cccagaacct ttccctggga ggaaactaat 17640
ctgcttagcc cagattggac gcagttctgc acagcacttt tccgaatgcc tctgaaatga 17700
gtcctcactg acagaacggg cccactctgg gggaactgag ggctctcttg gtcctgcact 17760
gctctttgcc atacagatct gtctgcccag gatttttctt gggtgtgtag gaggctgaga 17820
gagctcccct ttcttctcat ggctaaatcc cttggtcttt ccagccctcc tgggggttag 17880
aagggagagg gaaaaaaaaa aagactgaac ttgttgttgt tgtttttgtt gttgttgttg 17940
tttgcctgtt ttctatgttg tcttgtgggg agagggtata agattgattg acagagtggc 18000
acacttcccc tgcaaattca tcatttgaat ttctcaggta agatgttcac atttctctgt 18060
taagatgctc caatttctct ggttaagatt tctctggtaa gatgctcatg aattggtgga 18120
ggtgttggcg ggatgtggga agtgtgcctg ctctttctga gttttggggg aagttgcctt 18180
aattctctgc atgactttct ttgctccttt gggcttcatt tctgtgcaat gtagtctgac 18240
atgaatactg ctcagggagg tgttgcttcc cactgcccac gccactggaa accagtagcc 18300
caggtttact cgagtcctcc ttttgaggaa cccaaattct ttcatttctt ttatgtgaga 18360
tctgcccaaa atgccattgg caagctgtac tgggttgaat agtgtccttc ctcctcccaa 18420
atgtatgtct actccaaacc acaggatact accttatttg ggaatagggc ttttgcaggt 18480
gtaaccatta atagttatga tgaggttata ctagattaga atgggcccta gatcctatga 18540

ctggtatcct tacaagaagg ccatgtgatg acaaagacaa agaatggagt gaggcaccca 18600
aggaactcca aggattgcta ggaaacacca gaagcttgga ggaaggcatg gaacagattc 18660
tcctctcgga cctctagaag gaatcagtcc tgctgatacc ttgattttgg acttctagcc 18720
tccagacctg ttggggagaa tacatttcta ctgttttaag ctaccacgtt tgtggcgatt 18780
tgtcacagca gccataggaa actaatacat acaacctgca caatgcctac tccagcattc 18840
catagcaagt caagggcctc acaattatgt ccaaaggact gatagaagag cgacctctgt 18900
gctacttgtc cctcaggacg ctgacccaca gctctcaagg caggagtagg ccagagctca 18960
ttcaacaact ttgttatata ggggttccaa ttgtaaacct tttgaattcc tgtttgcaag 19020
tagatgaggg ttgaaaaata aatggccact ttctctaagc cacatacccc aatctgtttt 19080
gttacttcat tacagctgtt ataatggcct cctcttctat cttccaatct ccatagccct 19140
ggttccttga tagttctttt ttttttttt tctttttttg aggcggagtc tcgcactgtc 19200
gcctgggctg gagtgcagtg gcacgatctc ggctcactgc cacctctgcc tcccaggttc 19260
aagcaagtct cctgcctcag ccacctgagt agctgggatt acaggcacct gccaccatgc 19320
ctggccaatt ttttgtactt ttagcagagg tggggtttca ccatgttggc caggctggtc 19380
ttgaactcct gacctcgtga tccacccacc tcagcctctc aaagtgcggg gattacaggc 19440
atgagctacc gcgcctggcc agatagttct taaacaactg cccagaagtt ccagcctagg 19500
caggggcagc catgaactgc attgctcatt tctgcttttt gacctttcg atggctgaac 19560
tctaggccat ggaaaacaag gacccactgt atagttaaga gtcattttgt gactagggag 19620
acaaaaaagg gcctattctc caaatcccct ttcctctgg agttcctcgg tgccttaaag 19680
cttgtcctga gctacaggtg tgttacctgc ttatcccaaa atgcaggcat gttacctgct 19740
ttcctctgca aagagaggca ggcctggctg gggcacagct gaagatgtca aggccaacct 19800
aagggcagcc aagctatggc tgtctgtgac aagaggagag cagcggtgat gggagggtag 19860
gaggcattga gttcatgtcc gggtttgcct cctaccctcc tatcactgct tgatgatcct 19920
atcactgtct tgatgagttc aagacagaag tttgcctcat cattgccaca ataaaatcac 19980
caataacaga agtgtgaaag cagcgatgtg agtggaagcc catatataca caggggtaa 20040
tagagcagca tgattaaata tgtggccttg ttatcagaca ggctgatttg gagtcccagc 20100
tacttgttgg tgacctgaac tagaggaagt tatctaacct ttcattttac tcatttacat 20160
aacatggcta ataatagcac ctaccttata gggttattgt gaggattgaa tacaattatg 20220
caatataaaa cgtttagcat agtgcctagt ctaaattcct caccaggggt atgatgtact 20280
agtttttagt taagtaatta gtatcctgga catgtcacag ccatttgacc tatctgggcc 20340
agcgttttgc tcaggttccc ccagcagtaa ttgtattccc tccccaatcc cgggattagc 20400
ttttaggaag aaacagttga tctaaagata gaaagtcaga gtactgtctg gaggaaggta 20460
gagggaaatg tcattatctg ggttttcttt gatgatgtca gggaacatga caggctgctc 20520
ccaaagacag agcagcccca ggacagggaa gaaggtgacc ttgaggttga ctcctctgca 20580
tcccgatgtg gacgttatgg acttgttttg gagatgaagg gaaagaaaga tggaatgtag 20640
aaagtgaagg agaataaaag aagtgggagg aagaagggct gggaggagga tgggcaaagt 20700
ctttctggtc tcaaggataa ttacatgtga aatcacttgc cagtgggact ctggggctgg 20760
agcagctaca ataattacag tacaggctgc agagggctct tgggcatgtc ttggagcagc 20820
ctgtaggcag tactgaggcc tctctcacta gacccatctc ccagatcaca tagtacacac 20880
accttccacc cccgggcctg ttaatgatca aaaagcttaa acagaacaat tacagcttca 20940
gagtggaacc atatctctgg gctcctgtga tgaaaaccac aagcctgtca ggctggggct 21000
gcttcacatg gagggccctg ctcttaatgg ccaagtgatc tggagcaaga cccgtgactc 21060
tcccatagtg ctgtggatgg tgctgcctct ccccacgcat ccccagaaga ggaagttcag 21120
taactaagga attaactatt ctccagcctg attctgcttt tcccaatcag ggctttatac 21180
ctttcttttt catccctata tttggagatg agtcaccctt gccttcattt tacctaagca 21240
aggcagtttc ctgtaaccta atgaagtgcc aaacaatact gtgatttatt tagtacttac 21300
tgtgtgccag gaattccagc aggtgttgga catttatgat gtatgatcct tacactaagc 21360
ctgcaatggt gcaaccccag ccctgaccac tctgtgcttc ccttttcaca acacagcttg 21420
tcactaaatc caagtcagga attccaggtt aggcttgagt tgtgcagagc ccttaactga 21480
aatttgccat ggttgaggca tgattgcaat cactgacaac tcctcccggc tctacacacc 21540
tacttgtcat attcacgccc tgatcacggc cccactcgca tctcttccca ctttagaagt 21600
tctttcctat agaacacgtt gctgctgccc tgttctggtc actgatcagc cctggcctaa 21660
ccactggcta agctttgtgc ttgcacatag ctggttgaat cgtatgtatt gctgtttgtg 21720
tacatcaaaa atataataat aatatcggca attttatgtg tttcattcaa catgagggac 21780
ccagcattct taccttgtcg ctttgtaaac cctgctgctc tcaaatctcc actagctgtt 21840
tcctgagcag aaggagataa aaggctggct cacaccccca tgtttttact ggtcacagtt 21900
actgccacca tccaaggctg aagagacttc ctttgtgtta gggctaaaac cttagtcatt 21960

gtatctaaat gtcttctgta ttcctttcct caaaagaaaa aagtaccctc ttctgccaac 22020
cctctcccat gccaactaaa caagcaagca agcaaacaac aaagaaaagg tgatattaca 22080
gatgctgctc agcctatgat ggggttacat cctgataaac ccatcacaag ggatgtaatt 22140
ccattgcaag ttacaaatac cataagtcaa aaatgtattt atttcatata acccacagaa 22200
cgtgatagct tagcttagcc tacttgatca tgttcagaag acttatattc gtctacaagt 22260
ggacaaaaac atataaaaca aagcctattt taaaataagg tgttgaatat ctcatataat 22320
ttattgaata ttgtactgaa agtgaaaaat agaatggttt tctggatact caaagtatag 22380
tttctactga atgcatatca cttttgcacc atcataaact tcaaaaattg tcggtcgaac 22440
cttcctgagt caggaatcct gtctgtacag ggtataaagg aggaaagcat cagctttgga 22500
ggcaggtgga cctgtgtttg aaccctgatt ctgctagagc ttgacaatgc atattcgttt 22560
tctattgcat aactaattac tacaaacaac acatttattt ctcagttttc atgaatcatg 22620
agtccaggca caatttagct gcagttaagg tgttagctgg ggctgctgtc ttatctgaag 22680
catgggggtg ggggtgtgga ttccaaggtc aggtggttgt tggcaaaatt aattttcttg 22740
cagctataga actcatggct tgcttcttca aggacacggg gagagagaat ctctcacatc 22800
ttttaaaggg ttcacctgat taggtcaggt ccactcagga cagtttccct taaagtcaag 22860
gcttaatagt caactgatta gggaccctaa ttatatctgc aaaatacctt caccattgcc 22920
atgtaacata atcatggcaa ataatcacag gtcccaaatg ttcacaggtc ccactcacac 22980
ttgagggagg ggattatata gggcatgttc ttgcggagag aaggaatctt acagccacat 23040
tggaatctgt cttccatgct atttgacctc aggcaaattg actaatctct tgaaggttca 23100
atttccttac ctggaataaa aggacaataa gatcagccat ataaggctat gacaaagact 23160
aaatgagata gaataggctg gaaaagtctt gcagatagca gacacaagta tataacaatt 23220
tccctcctac tgttcctttt gtttttcacc tatcctgcag tctctgtcac ttcaaatacc 23280
atagaaaacc tttccaagca gcccaaatca tgcccccaaa tagtcacgtc tcattattca 23340
tagcagttat gttccataaa gttagcacaa actccgaatg agtgaatcct aaagcgttgc 23400
tcctggagga aatacaggct gctggtcaca atatttttat caactgatca atatatacct 23460
tgtcttatgt gtgtttctgc ttcaagacac tttatttaat atatacgttg attcattaac 23520
tctgaactct ctaggcaaca gcattataac tcctgccttc acaaagctta tctaacacac 23580
acatttcctc ctcaggcaca tcccagcctt cttgcactta ggattcagca gtatgcttaa 23640
gggccatttt caacagcaaa ctcatcagcg caaacacaaa catgtgaaaa acgtagcact 23700
aaagagactg caaaaaggac actggcttac agcatggaag ctggaaggag aaggcagaga 23760
atcaccttgt tccacttcag ctatgaatat gcagtcaggc cacccagtca ttcaaatttt 23820
ataaatatac tctaatatat atataaatac caggcagggt tatttttttc ctcaagtcat 23880
ttttctaatt ttttttaaat gaatagatag aagagctgaa gtaagggtca ggagcaagag 23940
ctctgcttcc ttttcccttg ctgggcttcg ttagagagcc atcatctcct caatatgtct 24000
cccaactctt ctaggcattg gatgagtttg ctgcagatac gaaacccaac tttgccagtc 24060
acttcatact aacaggtgaa atgtagtgga ggagcctttt gaagacaggg actcagcccc 24120
ccattagcct cattgcagac ctagattcct gccaaaatta atttggctgg aacttcccag 24180
ccatggcatt gtcgacatta cacatcttcc actgtaatgt caattaccat tttattcagc 24240
cgaatgctgg agagttaatg ttcaagtggt tagagctggc tacgggtggg ctgaacaaga 24300
tgtcttttcc ttcatttccc ctgcctgtgg tgaaggattg taaccagccc tggctggcag 24360
cactttgaag ctcacccaga gtgctcctgg ggacatcttc tacagagcct atcatttgga 24420
catgctgtct tctgggcctg tcttccttcc ttccttcttc cctccctccc tccctctttt 24480
ccttccttcc ttccttcctt ctttccttcc atctgcttta aaaccagctg ccttgagtgc 24540
ttgtcttggc gcccctcatt agtgccattg caatcatccc tcctgcctac cctgctaacc 24600
acagcttgtt agtccacaac agcaacagct gtgtgctggg gtgcagcagc tggagggcca 24660
aaggtagggc tgggggacag ggtgttggga tggttttctg gggcagatga gtttatacgt 24720
ttctttcatg tccccttcct cccacataga cttttatttc cccaaaggaa aacagaaaac 24780
aatgatctgt ttgacagtgt tgctatcatt gggcatcaaa cctatcatct aaggggaatc 24840
cccctgtata atcagtcagc caaatggagc aggaccctgt gttttgtagc tgatacaaca 24900
gggcagcatc tctagtgagg gggccagggc ttctatttcc ttcattaaaa aatgaaacag 24960
cagacctgat tccatattta gagattacac ttagttgcca ctgtgggtgt gcaggcacca 25020
accaaaccca gttggcaccg ttgtcttttc tctgcaatga tgtattgaat ttaataatgg 25080
aggtatatga aattcagagt gattggaact gaaggtttag gggctttgtg taaaattgat 25140
atgtaaggga tttggaagta ggtgagggat tcttccccaa tacttattca attttggagt 25200
caaataacca agcatttaca aatagccaaa aaagaaattg aaagagggtt taatccaata 25260
aattttcatg cctcatatga accacatctt ataataagaa ttatgctttt tcatttcata 25320
ctcagttaac aaatatgatt tgtgagcacc tggtaagttc agggcactag gctgaaaggg 25380

gttaccaaat gtcttcattt aacaaagtcc agctgagctc ttacaggtac cagaactgtg 25440
cctgggctgt catatgaaga tgaatgtaag agtgtgtcag gccttcaaga gcttacagtg 25500
tgtcaggaga catcaaacaa gtgagccaat aaaatgatac tgccatttta gaaatagcct 25560
gaaattcatg gagttcacag tcttgttagg aaagtgaaac ataaacctat aagcattaaa 25620
aaataactgt tgaagacagt aacggaagaa tgcaactggc aactgaatga tataggttgt 25680
gatgactgtt aaatatcatg aaaagagacc atgatgagct gaggcactcc aagagacttc 25740
tttttggaga tatgtttgga gccaaatctt gaagatttaa ttgctttttt ctttttttt 25800
tttttaggtg gagtctcgct ctgttgccca ggctggaagt gcagtggcat gatctctgct 25860
cattgcaacc tctgcctcca ggttcaagcg attctcctgc ctcggcctcc tgagtagctg 25920
ggattacagg cgtgtgccac catacccagc tgatttttgt atttctagta gagatggggt 25980
tttgccctgt tggccaagct ggtctcaaac tcctgacctc aagtgatcta ctcgccttgg 26040
ccttccaaag tgctgggatt acaggcatga gcactgtgcc tggccttttt ttttttttt 26100
ttaaaaaaaa aaaaaaaaaa aacaggaagt tttcgttagt ttttttgttt gttttacttc 26160
ccataaaaac tctttgtgtc acatggaggt gaatggaaag agaggctgtg gcaacagacg 26220
ggagactttt ctgatatcag aacccagtcc catagaccag aatgtatgct ttcaatccac 26280
gttgtctggg tccatcctat tgagtgccct gcccccacag cggggtatgg agaagagtca 26340
gacacagccc cagtcctcac gtagctcaca atccagtgga ggagacggac tcagaaacag 26400
atagagatga agccatgaga tcagtactgt ccgaggccat ggccacggtt ttgtgggaac 26460
ccacgagagg gaatgactaa ctgtggggaa gaagagggag aggaccaaaa tgcaggggaa 26520
gtgctcacag aggataagta agcagtgagg tgccatgaaa tgagtataca cctgacagcc 26580
gtgtaacagc tcagagcctg ggtagagggg aatagagctg ctggttctct ggggggaaga 26640
gaggggtatg ggattctgga acagaagcac caaaaccagc aggttattgg agctgttagt 26700
gctcagatca gcaatgggtg cacaaccaaa ccattctcct agggatgagt tctttcctgt 26760
ggatgagggc ttctcagcct ggcttctccc gagaattacc cgggaagctt gaaaagtact 26820
gatgcctgga acctacctcc agagagttgg atttcattgt gttgacgtgg ggctgggata 26880
tcagtatatt gtttaagcac tccaggtgat tctgatacgt agctgtgatt gagaaccctt 26940
gccctaagct atccatctgc actccagggg tgctcccagg cccatctgtt tgtaaatgga 27000
caggtgtctt gaggtaacaa atgtgccaag gctctggagc caagcacgcc tggctcctta 27060
gtgcctactt agtgacctca ggcaagttac taaatggctt aaactttaca aatccttaat 27120
ttgtaaaatg tgggcaatga tagtacctcc tcacaggatt attacgaggt ttacacggaa 27180
tactctcagc tcataataag cacttgcaca ggcctcatgg gctaggccct caaaacttaa 27240
cgcatctaca ggcaacagcc atatgaaagg aattttatac caccaagtca aaaaatctgt 27300
gagcactgct cagaagcaaa agcctgtctc caacagcgct catttaaggg gtgggcgagc 27360
tacagagaga agaatgagcc cccacagggt aagctgggga aagctgggga cagaatgaga 27420
ctcaggaaat cacttgaata ttgattatat ttgtgctcaa taataaaata acgaaatgag 27480
tacagcccta gacctaaaca ttgtgggtga ggcaaaggca atgcgttaat tttgcatcca 27540
ctgaggaaaa actctaaaac ggtgacttct tttttaaggg accagaagaa tctagattat 27600
atttagtcta agtcaataca tacgacagaa ccttgccctc tagacttgat aagaaagaag 27660
taaaataaga gaaagaataa aaaacccttc caccaaaata ctaacattca gataatgact 27720
ttttagttag gtctcctgga gaggaggttc cctcagaaat gaatagattt ctcttctagt 27780
gcaatcatca aaaggtaatg catggactta agtgtgatcc ccaagagaaa atcaatgacc 27840
tttctgtgtt tgcctttgag aaaatcagcc agtctatggt taaattagac atatttttc 27900
tccttggtca agattagtgg gaccaagaat gcagtcttac actccttcta gcaaagaatt 27960
acctgatgcc ttatttcaca caaatttgca aagttgtatg gacgttgtat cttattttaa 28020
ggagaactgg tgatcaaatg atgactattt caatagtggt tcatttacac caccaccctc 28080
accccacatc ctgctttcac ctgaatctga acgatcatag tcagtctgag attctgaagg 28140
tttgaaattc cttttctgag ctctgcaaga acagcatctc ccaagagagc tcagggcaga 28200
ctgtctggga gagattggaa acctgtcttt tgcagtaaca tgaattggtt gaatggtcac 28260
cctccatatc aggcctgctt ctcccattgg gtttctgatc agcccaactt gggtctcacc 28320
cttctgattt ctctctcctg gctcacatgg ggctgcactg gccattaggt gccaggcttg 28380
gctccgtgga acccattggc cagctgggct ctgtggagcc ctaaggcagg gctctggtca 28440
ctggtgagag ggaggccatt ggagtcactg gggtggacct acagacccta gggttaacag 28500
ctaggtgggt gtcctcttca gagaaacggg ttacaaagtg aaagaaagtt acactgtgag 28560
gtcagccagg gaggaagaca gagagctgat ataagatagg tactgattcc ctggggatgt 28620
gaaaggaggg taatattcct aaaatgatag catttagctt ccagtataca ttaattgatt 28680
cctgatattc attaaaacta aacgctattt ccttgatgtc tcatccaaag ccgcaccact 28740
cttcccacta agtctgaggg gagcttgttt tgttgacaag tgtaagaggt tgaagaggga 28800

cccatgaact cttttgtcct actgaagaga tccacagatg gaaacaaatg ctcctaccac 28860
atttatgaac tgctgctttg cagtcccgct tctgctatca tgcacaggaa ctgactaagc 28920
tccaaagcca gaggatgtaa atctccctgt aataaatgta agtcatttat tagctacata 28980
cacttcagca agtcacctaa cctgcaaatt tcaagcatgt gaatcttgga tctttcatgt 29040
gctagctgtg agactttgag aaatgtattt aatgtctctt tgcttccttt tctacccaca 29100
caatgggtat aataatgtct accatatatc tttgcagcaa ggtctaaatg gggtgataca 29160
tgctgaatac atttccaaca gagtctgtgc aatgataagc tctttccaaa tgttagttaa 29220
agctaaccaa ctaacccacc aacaaaccaa cctcttagcc aggactgatg gaaggagtct 29280
gtgagagaat gcatttaaaa cacttggcac catgcctgac aagagtaagt actcgataaa 29340
tcagttattg ttattatcgc atcggtatta tgaccattat cctcttctct ataggcttca 29400
ggttttcctg tcttttttatc acagcagtat tccagcagaa gcctttgatt taactaagtc 29460
tctactgtgt gtgtggctag atgctataaa gcatccagag aagtgagaat ttggtcctgc 29520
ttttaagtag cttatagtct aattaggggg aagtaatcag atagaaagga aactaacaat 29580
atgcaaaagg aaactcatag tttgtggtaa atgccaggtg ctgctgatag tggcttcaga 29640
gagatctcat agatgctata ggaggtcaaa ggagaagcgt gcagcttgag ctaagttttc 29700
agggaaaagg gtgaaagaat tagtcattaa tgtacaccta cattacctgc cagactccat 29760
tcaaaaatat tcttaccaaa tcatcacaat accttgttgg taggtactat tactatttta 29820
cagaggagga aagtgaggca aagacacatt aaataatttt cccagaatcc caaggtgtga 29880
ggtggagcaa ggacacaaat ccatggctct aagtccctcc tagtatatcc tgcaaacaca 29940
tctggaatta atgcagagag gaaggggaga ggcagtgttc tgcaggagtt cagagccatg 30000
ataacccttc ttgtgtggct tttggtaagt tattttacct cttaccctct gtttccccat 30060
ctgttcaatg aaggttgtat atacacacat tatatggccg ctgtaagtgt gcagtgatat 30120
gatgcatggg gactcagttc atgaggcagt gtgaattctg aaggtatcac aatgggacag 30180
gtgtttttttt ctccactcat tttctccgaa agtcttttgt tttgttgccc tccctctttg 30240
gggcatatgc tttcagctca taccttaatg acatcagaat ctgcaatttc ctggcaactt 30300
ttgtggttaa aattattctg cccttccatt ttaaagcact aatagcaaag gtattaggtg 30360
caaaatgatg ataaaaataa ttgcaatttt taccattaaa agtcatggca aaaccacaat 30420
tactttggca ccagctgaat attttgaaac tccctactct gatgttaacc aagttcatga 30480
ttcaaagaac ttgcagaggg gtaggggaat ttcaagggaa agggggagat gcctggggtt 30540
gtcacacact ctgtctttca tcctctattg acatgttggt tatttggaga tggtattcag 30600
ttccactata gcccctcagt cactgtagac cctctcaaag gggcaatcat gtttccctta 30660
ggtcaggtcc attcatctaa cccctctccc gggggcatca ccttgtttgt tccagcagct 30720
gtctggccaa actcacacct cctcctcacc ctctagccct tatgatctgc tttggggagc 30780
catgggaacc cctagtttcc tctttcatac ccactgagat tcacaagtaa ctaaggtcaa 30840
ggcggggctt cattgccttt ctgcagatac cttacgctac tgttcctcct cgcctggctg 30900
gctccacact ccagcagacc ttctgctggg cgagaagctg caggcctgaa tctctgtgtt 30960
ctcatatggc cccaactctt gggattacac tagctcttgt aagaactcaa tgctctgctc 31020
tgctcatttt gatgccatca aagagggctt gcaagttacc agctgggagt gaacaccagt 31080
gtcctctttt tagaggtacc cctaatcttt ctgaacaatt ttgctggcac cccttcactt 31140
ggctttgccg gggtaagagg gggcacttct ctcctttccc tcatgaaagg agggagagaa 31200
gccaaaaatc tccctactag tcaacaactc aggcacccct ccttctctcc tctattttat 31260
agactgggaa gggagtgatg gttgttggag gtggcagagc cagttcagct gccttttgtg 31320
aagtcctgaa ggaggtgtct atcctcaact gctggcttct gtccttaagc ctggggagaa 31380
ttaagtcctc tttgcctcag tttggcactc caattgccaa cattgggaca gcaggaaaag 31440
ttccatccaa catcccatta aatatgtaat gtgtattagc acagcgcctg gcactgggca 31500
ggtattttct aagtgatagc caatgcgaag cctactttat tattttcctc tttgcttaac 31560
ctacaaggtg tctaagacca tttgtttgtc cacacatagt aagataaaca gcactgagac 31620
tgtggtcctt tctgccctgt gtccttatcc cacctgggaa tctggaaagc caagcctaga 31680
cacactcgtt ccacaaatgt ttactgaagc ttgttctatt caaagcactg tacagctaca 31740
aagaccatct tttctgaact ccaaaccagg ccacatggtt ggaataactt caagtatgga 31800
gaccaagaga aaaggtggtt gttgtcagca aagctctgag tccacacctt ccaggaactt 31860
atagttgatg caatggtggg agaagtctga acctggattc aatctgcttg attccgatga 31920
atggtgcagt aggcagagcc atgagttcag agcaggaaga aaccactggt tcaaagaagc 31980
atctgtcaca tcgaagctgc tttatagtct gttgggaagc atgcataata atttattctt 32040
tctttctttc ctttggtcaa caaagatttc ttgagtccct actatgtgcc aggtactctt 32100
ctaggtactg aagatgcagc agtgaacaaa gaagatacaa tccctgccca gcggagctta 32160
cattctagtt atcgaaagtc cctttctcag tggctgctct ctttatttga gaaaccatgg 32220

gctgttctcc tcccatccta gggctgctgg ctccacagag gcacacagtc catcaggatg 32280
ctctgccagc cacccaccca ctcaagacca agggttacgc tgtcagtgtg agcagggaca 32340
ctcccgtctc tgctacctcc tttctcctga aaacaagatc tcagggaaca tctgccatcc 32400
attttccctc cctggggagt gacaggaaag gtgtatggag gagattgagc ggagtgatgg 32460
attgaggcac tgtgaaagtg aatcattgcc tgacatggga atgaggagac ttgcttaaag 32520
gacaagccat gctaagtcat ccatcgttct cccctaagga ggtgaattga agttcccatt 32580
tttcccaggg agccaaatta acaaggtgct gggagatttc caaattagaa aaaaaaaaaa 32640
aaaaaggcac caccagctct caaatcagag aggctgttga gttgtttttt ggagcagatc 32700
attgtatttg gcatctaacc ttgaaataga ggagaaagca tggaatttct gctgaaaact 32760
catccttctc tgagcaggtg gtacaaataa gcatcgttgt gttctcagag gcaggaacca 32820
catttgcacc ttgataccaa ctacctcaat aaccacagtg ctgaattttc acaaattgcg 32880
aattaggaaa ttgttgctca ttttacaatt tggtttccct caggattcct tttaagtagc 32940
cagctacccc agtacttttg aaatatgact tgcttataaa aatttgatag gcttggcacg 33000
gtggctcaca cctgtaatcc cagcactttg ggaggccgat gtggggtgga tcacgaggtc 33060
aggagttcaa gaccaacatg gtgaaaccct gtccctacta aaaatacaaa aactagccag 33120
gcatggtggc acatgcctgt aattccagct gctcgggagg ccaggcagct aggcaggaga 33180
atcacttgaa cccaggagat ggaggttgca gtgagccaag atcatgccac tgcactccat 33240
cctgggtgac agagcaagac ttcatctcaa aaaaaaaaaa aagatatata aacaagtttt 33300
tataatattc tcaatatgaa ctagtagaaa aaaagcatgt gtttttaggt cttagaggcc 33360
tggttcccag ttttatctct gactctaatg aggtatagta ttacctacat tgattagccc 33420
ttctatactt cataggagat gctccaagac tgctagcttt cttcattcaa taaagagaga 33480
tataacagga tgggccttaa aagtagcatg catttcttct ttcattcact cattcaaaat 33540
attttcatgc gtgaaaatgc caaggatgtt tggtcaacca actcttccca gaccctggct 33600
gtgagcctgg cttagaacaa ttccatttta atggtccatg ccctcaggca cttgtattct 33660
agtagaagag caaggtaaga aaacagctta aaaagttaaa cagttttagg ttgagatggg 33720
tgttgtgaga aaaataagca ggatgctttg aacctatgca ggtaggaagg tctggaaagg 33780
cctctctgat atggtgatgg ttaaagcaaa accaaaaaga ccaagaacac atggaacaca 33840
tgaagggctg gaagaacagt gttttatggg gaaggactag tacacacaaa ggctgcaaag 33900
gcgagtgggc tcattatgtt ctagaacatg ccaaaaagcg ggtgcagctg gagagggagt 33960
aagatggcac aaaaggtgag tgaggtggac aggagcctta tcacgcaggc ttacacaggc 34020
tctcagaagc cctgcgtgtt ggtttcttgg gactaccgta acaaagctcc acatactggg 34080
tggcgtaaaa caacaaaaat gtattgcctc acagttctgg aggccagaat tccaaaatcg 34140
ggtgctggca gggctgcgct ccctccaaaa cctgtagagg agaatccttc cttgcctgtc 34200
cctagcttcc agtgggttgc tagcaatcct gggctgggtg actccagctc tgccttggtt 34260
gtcacagggc gttgtctttg tgtgtctctg acttcacata gccctcttct tcttcttttt 34320
gtgtgtgtct gtgtgtgtcc actctgaggc acagaagttt ttatttattt atttattcat 34380
ttatttattt cattgataaa cataatagtt atgcatagtt ttggggtaca tgagatattg 34440
gatacatgtg tacagtgtgt gataatcaaa tcagggtgat tggaatatcc attcacctcc 34500
aaacattttc tcatttcttt gattggggac attataattc ttctagctat tttgaaatat 34560
acaatagatt attgtttact ataatttccc tgctgtacta tcgaatacta gaacttattc 34620
cttctgttga gggtgtactt ttgcacccat taaccaactt ttctttatgt cctccttccc 34680
acttccctta ccagcctctg gtaaccacca atctactctc taccaccatg aaatcaactt 34740
ttttttttta tagctctcat atatgagtga gactatgcag tgtttgtctt ctgtgcctgg 34800
cttatttcac tcaacataat gacctccagt tctgtccatg ctgctgcaaa tgacaggatc 34860
ttatttattt ttttatggct aaatggtatt ccattttgta tgtatatcat atcttcttta 34920
tccattcatc cactgatgca tatttaggtt gattccatat cttggctatt gtgaatagtg 34980
ctccaataac catggaagtg aaaatatctc ttcaacatac tgatttcctt tcttttggat 35040
atatacccag tggtaggatt gctagatcat atggcagttc taactttaga ttttaaagga 35100
acctccatac tttttttcca tggtggctgt attacttaca ttcccaccaa cagcatatgg 35160
tcatctcctt tctccacatc cttgccagaa tttgttatat tttgtctttt tgataatagc 35220
cattctgact ggggtaagat gatatatcac tgtagttttg atttgcattt cccttataat 35280
tagtgatgtt gagcattttt ttatatacct gttggccatt tatatgtctt cttttgagaa 35340
atgtctattc aggtcttctg cccatttta agtggattat ttgtttttt gctactgagt 35400
tcttcgagtt tcttatatat tctgatacac agccatcttc ttatgaggac tccagttata 35460
tacgattaga gaggtccacc ctttttcaga atgaaattat agcttaacta attacatctg 35520
tagtaactct atttccaagt aaggtcacat tctgaggtac aagggtttag gacttcaaca 35580
tatgaattcc agtgggacac agctcaacac atgacaccat ggtagggaac tttattctac 35640

ttgcaagttc tgagtgtctt acgcaggtag atggactggt gtgatgtatg ctttaaagac 35700
cgctgtgtga agatggcctt agggtgatga ggatggaagt tggagactaa taaaggacta 35760
agaaaatgct aagaaaatcc aggtgagagg tgatgatggc agaactaagg tgatagcagt 35820
agagagaaga gaagtggatg gagattagac atcttttgca gaacgaatga caaaataccc 35880
ctatggattg gacatgggat gaggaaaagg aaggacttga gggtggtgtc taggcttttt 35940
actttaatcg tgaagggaag ctggtgccat ttaccttgtt cggacaaacc tggagaggat 36000
caggttaggg actgcgagtg gtatggacgg caaaggaatg ggaagaatgc agggattaaa 36060
aattggaaat ccccctcccc agtcaacaat atcttacttt tatctgaaaa atactaagta 36120
aaaaagcatc cttttgttgg aaagctcaat ccttgttaaa atgaagacat ctctgggaga 36180
ggaaacatag tgagcacctt tcccaaaagc agccactgat ttggagatga gacagagtag 36240
catacaggac atcagagaga acatgctcag gacagaaaga gcaatgtagg acaaggcagt 36300
gtcttggcat cacagtcttt cctccgactg gctgtgagca agtgctcaat ttaattccat 36360
ctcagtgctg ggtcaggaca agtgcccaaa agcaaattga caaaagtacc agcatgatgg 36420
agttagaagg tagcaagttc cctccacaga gcccagctgg aaaggaagat agaggggaag 36480
ttgacccctg gggatgggga atagggtgag aggagaacat gaaactgaga aaagggcttt 36540
gagtgaaatc taggctaaaa gctaaggttt ctttagaaac ccaccattga cccaacatga 36600
ccagggcttt ctcttgactt gattattttt gataccccat cttcttctgt attcctggaa 36660
ctagctctcc caagccccag aattgtgctt ctatcagagc tgggttttca tcagagtctc 36720
ccctttatcc tgtatctctg ttgccctatt ttgtttgaat tcctgccagg tcagctgaat 36780
ttgggcattt ggggtgaaaa accatcaagt gtggcatcct ggctttggca cctggcacag 36840
tgtgacccca ctggtctctc cctcacattt gctgtggtcc gtgcacggaa tttgtcaaaa 36900
gacctcctca gtatcagctt tcctgcagcc tcaatgcacc ttgttctgaa taggatatta 36960
ccccccaaga gtatattagg gcattttcct atgccagaag gggtccttag gcctcttgca 37020
gtttttttctg ggtgacagtg aaggaggagg tggctgcaga gcttactgcc tgtggactga 37080
ccaccccagg gcctggtgtc aggaccattt gtccagcctg ttgagtgaag gtcattctgc 37140
ctaaactgta agcacaagag agagttcagc atcatttgca tcctatttta ttgtctttct 37200
tctcttttct ttcaaggcct cattttttt ggcttgaaca aatggtaaag gccattttat 37260
tacaggtacc aagccaaact ttccttggtt ttgtggccat cctgctgggg aaggaagtac 37320
tcctttactt taaataactt taaaaacatc tgtttggtct caggggctgc agctggaaag 37380
attttctaac taatacttgt tttatggggg tgtttttggg ggggtttatt gagtgtcaaa 37440
cctggcagta aattagaatc agaagacaac agttagtgat aagcagagaa gccaaggatg 37500
ttaccatagg caggcagcag agagagggga attggtggct ggccccccaa aaacagattt 37560
gaagatctcc ttctgtcatg tagtgaatcc ccaagtgcct agggtgggct gtgattactt 37620
gagctcctgt ctccactgtc tcagctcact tgccttgggg tggacacaca acacacattt 37680
gctcatagca tcaggtattc aggagcaaag agctgaattt atctggttaa tttagatacc 37740
cctacccccct cttttaacac cagattgcca ggatcatgac ctcaaaaggc taccctgaaa 37800
tgcaattgac aaatgggatg aaagatttcc cgtttcatcc acatttgcct cctgagctac 37860
ttacagcagc aggtcaccgc agccagagcc cacctgcttg cccaccatgc ccgcacacag 37920
acaatgctgc ttctgtggct ggaggtcgga acacctcagc actatctcag tttggctgca 37980
gatcctctgt gtgcttggta aacaggtttc ctcatctgta aaatgaattg gctcttccac 38040
aacttttta aaagcactaa catattagga ctctcactaa atactcaaat gctaaactca 38100
aatactaaaa gagtgcaaag ggatgggctc ccaaatatta cagtgaaggc tgcagcattt 38160
tctgaccttg ctgctttttc tggtgagtgg cttttatttc ttagtttggt ttcttctctc 38220
ccattctaat caagcaagaa gtgaccacca aaaggggcac tcaccaaacc agaacaagct 38280
agttctttca tctttaattc attgcaacca aacagatgcc acagaaagag ccaagggctc 38340
caggctttag ctccagcctt gccattaact acatatgtaa gtcagccatg ctggtctgca 38400
ggttcttgct ttgcatgatc aagggacaac ttggaaggtc tccaatcact ctattccccc 38460
agatggaaat gtattcactt atttcctgga gatgtctgtc ctcctcccag ttaaagacag 38520
accttgaccc acctccactt ccttctctgt ggccctgtct tatctgtcct cttgttcttg 38580
cctcttcaat tgttctctca ccgtgtttgt cacttctgag ctatcactgt gatcccccctg 38640
attgtttttc taatgtccct gaacttcaac ctgattttca cgcatataca atgtcttcct 38700
aaacacttat agactctgac acattctgta actgacacat ttccctttat caaatgcaat 38760
ctaagaagct cacagtttct ctcagtttca acaagagaaa tcaggagcac ttgaattata 38820
caacttgaca ttattagggc tgatgtctga ttttgtcctg tctgcccctg tcatttctgt 38880
actacctttt acaaaacctc tcctatgacc tgtgtcctcc tccagctcca tttgagaaca 38940
cctgctgtat accctgtggg ctagctttta ttatgttcgc ctcaatgatg aagaaacagg 39000
cttggaagtt aaattatcta ccccaggccc acagcctgga acctaggatt ccaaccaaac 39060

cttgtctgat tctaaagcat agcagaggct ccatactctg cctccctctt ctacatcatt 39120
tcagtttctt cactttccca cctccaattc tcacccaaac tgaatgtctc acagtctctg 39180
tgcccccact ttgctccatc ccttggcctt ctgcagtcca agctccattc tgagatcatc 39240
caaggcttct cttctgtgtt gatccttggc cttcttggag tctctttctc ccatgttctc 39300
cacaacagag cattctcctg actgttttca ttctgcatct cactctttca tcagtatctt 39360
tttctctacc atgccccata aatttgggtg ctcctgaggg tcctgtcctt gtcccctgct 39420
ttcttgttgt acaacctcct tgatctactt catctactca agtttggtcc acaatttcta 39480
tattgtgaag attcaaatct gcatctctag ccatatatcc atttgcctgc taggcatttc 39540
tacctgaata ttttataggc atgccagtgg ctcttactct atggctctta ctctaagtct 39600
agactacagc agaaagcaat gctcttttta ttaaggcata gtgcctcttt cagaataatt 39660
tacagcatac aaccaggcct gctgtgcagc attacaattt gtcattaaaa ctccattcct 39720
cttgccagag taaatgagcc atttacagcc agggcgccaa gatggactgt tgttattttt 39780
tctgcctttg tattatgagt attcatggct ctcctcagac aagctcctgg ggattcccag 39840
tggagttgcc ttaacatgca ggtcaattag ccaggctcaa gggtagtttc ctggatattg 39900
gtatccccct tgcagaggac tgcaggaaag ctgaacagtg ttcccccaat gtgggtggtg 39960
atcctgagaa atatcatttg tatctgcatg tgctgtctca cacacactag ctcacatgtg 40020
cacacacacg tgcatgcaca ggacaaaacc aaacacaggg caacccagca tctgcccccc 40080
agccatcagc attgttacac ctttataggg ggcgggaaca ggttggtcag caggtgaacg 40140
tcaggtgagt tgagaaaagt tattaaatct taaatcctta aggaaagtta ttaaatctct 40200
tctaaaatgc atgcataggc gggctcagta actaacatgc aaatgtttag ggtctgaagc 40260
tcctaccgat aatctttcag atctcagaat tccagcccct tgtgctgttc tgggttgtct 40320
gacacagacg aagcagagaa cagtagaata aacagctcag taaacaattc attgagggaa 40380
agagagtgag aagattcact ggacagctag aggaggaaat actgctggtg actatggaag 40440
aaatttgccc taaggcctgc aggcaatagc ttggtcttat ttatcctggt gtcccaccct 40500
ctcctccaac acatactgcc ctggcaggta cgtagaagat gcgtgaaaat atcttttgaa 40560
ttgagctatg caaaaaatac tggattctgc cctccaagag tttactgttt agtttcacag 40620
aaagcacatg ccctcctttc tctgcctctt gaagactgac ctatctttca aggccactgg 40680
cccaattctg ttttctaagt aagaccactg agtcagtggt gacctctcct tctccctaac 40740
aaagtctgat ttacttgaat atacaactat ctccctcttg gcctgtgaat ttcttgtgtt 40800
agggaacata tctgatttat ccttatctct tccacagtac ctggtgtaaa atgcccaata 40860
aatgcattga aatattcatg aagcttacta aatgctctgc cttatgagcc atgaaatata 40920
aagtgcctta aactttgttt ttctcttatg taaaataagg ataataataa tgacacccct 40980
ataggattgc tgcaaggatt aagtgtgata atatatataa aactcttagc acaaacacct 41040
ggctcacagg aatagtagct actaccataa tggtaacttc gagggcaagt tttctcagag 41100
ttatttagcc ctccttcacc ctgtgtccag gagtgcagat cagaatggtc agattccagg 41160
acaccaagtt ttctgtggga gcttccctag gaatataact aaggaattta aatcaggttc 41220
agctcatgct gttacactct cttcctccac tcaggcattg ggtgtggctt ttccaagctt 41280
gagaagggtg tgatctgaga tgggcttggg tatagagggg aattatattt aggtctaccc 41340
tgtataggaa aaagtgcctt cccaaagtct ccctggccta aagtataaga gatatgtgtt 41400
gggatttaga cccagagccc aagccaataa tgggaccccc ttctcacatg tggctacctc 41460
ctgctatcac cacaacagct atcataccca taactacaac agaggccaat taacgtggtg 41520
ataattgaca aatgtcaaga catcctacat tgaggcacac tgtgcgtttt gcgtgagctt 41580
ttaaattggt agggaaggaa aacttttata cctacaccta tcatggaagg cagaaggtaa 41640
gagctaaaat aaaggtatgc caagaacaaa ggcaggaaag aagggtttta acaacttgag 41700
gcctgatcca ttgattagtg aagaggaaac atgttcaaaa accactctat aaccaccttc 41760
tccaagtttt ttataatttt gcttcttcgg atatcttctc atcatagtct taaatgccat 41820
caaattaact gaaaaatgct aaaaatgcaa ccactctaag agaatgggtt agatgggaga 41880
tggctttgtt aaagaagtcg gtcttaaagc aaaagtaggg ctttgtcatg gtagtatgga 41940
aggaaggaca tttttggtca agagaagaaa gtgcagggcc tgttgaggaa ggaatgagta 42000
gtaaaatatg gctagaacag ggtgcagagg ggaagaactt cagagaatga ccaaataaac 42060
aggctgaaag gtgtagacat tataggcaat aaagcaacca cagaggtttc taagccatag 42120
ggtgacatga tagatctgta ttctagaaaa gttagttttg cagcagttgt gtccattgaa 42180
agggacagga taagggagat agataagaag acatgctatg atgataacta gatttggata 42240
ccaagtggta tggtggaaag gaatgagaga acagggtcac agatgaatga ctgcccaatt 42300
tcaatccatc ataacaggat gtataggatt gcccttaagt aagatgggga atccaaaaac 42360
gaggaacaag tttgtaaggt tttgggggcc aatgatgaat tccatttggg acatgttgct 42420
ttggatatac caatgggaca ttcatgtgaa aatgatctcg gcaatcctat cctggaattc 42480

aggataggat cagaatgagg gacacagttt ataaggtaaa cagaatggag gtgatataga 42540
agataagggc atagatgagc ttaccaaagg ggagagttta gaatgaaaag aaaagaccaa 42600
aggctaagcc tgtgctattc tttctcctca caatacgctt cagacctggg cacaaaccat 42660
cagtgagtgt catgataaca ctactgtggg caaatccccc ctctataagg gcctgatttc 42720
ctcctctata aaatagaggg ttgaacaggg tggtccatat cctgttaatt gtgtttggag 42780
agcacacaac aaaccagcta ctatccaaag gggacatccc gaggcaggac taagcaaagg 42840
aaatccagca cagggaaaac actttctggt gctggtccca gttaggcagc gttcagttta 42900
acccatcacc atcaccatca gtagctttca gctgctactg accacactta taggaagaaa 42960
aacaattaga atggagagct aactctttgg aaatggtcaa agaacacggg tctacaaaac 43020
cgtcaataaa gcgctaagat gcctgggcgg ggtcaaaaag tctacctggg cggggtcaaa 43080
aagtctacct gctcagcata tggggcccag acatctgacc tttaccaact ccacaataac 43140
cacttcatct atggatccag tcttggtatc acctagtcgc tgttttcaag taacagaata 43200
tttggttctc aatggtaggt gactggaata cagcttactt tctcccaccc ctaccgccaa 43260
tcctttctgc ccccttatag tttaatttgc ttgtaaatta cttgggaata catttgggag 43320
ccattatagg gaaatagaag gcagacatga tgaacagaat gcagggtgtt ttttattact 43380
tcacattgtg ctcaacaatt aggaggaatt ctagaagccc ctcccagtgg ccaggaattg 43440
gtcatagcat gaataaactc aatataggtt gagtattcct tacccaaaat gcttgatacc 43500
agaagtgttt ttggattttg gatttttttt ttgaatattt gcattatata cttaccagtt 43560
cagcatccct aatccaaaac tgaaatctaa actgctccaa tgaacatttc ctttgagtgt 43620
catattggca ctcaaaaggt tccaattttg gagcattttc aattttgggt tttgggatta 43680
gggatactca accagtggta ggtttgggat gatatcagca tgctaaggtc aaagagacct 43740
agctgggaag ggtgggagga acatggaatt ttcattctct gggcacccct tgaacagtct 43800
tactattagg gccccaaatt tgttctaagt gtgtgtgtgt gtgtgtgtgt gtgtgtgaga 43860
gagagagaga gagagagaga gaattttctt tcttccttta tattctaagt tcctcaggac 43920
aaaattttgg gtttctttgt attctccctg cagctcctca tgtagttcta agcaaataaa 43980
ggaattcatt aggtccttga tttcagaagc ctcccagttc tctatgtagg aggaatctta 44040
gggtggcaag ataagttgag ggacttttct tcaagcacat ttcacaagta agagaaaatg 44100
ttgactgtgt atatctaaga atgggtgggg ctcaatgatg ccccccctaag ttactcttta 44160
ctattattga ttgattgatt gattgattga agaagcaatg ttttgattga ttgaagaagt 44220
aatgtttcca atggctacag cagactggag caaaagaaca aaatgaaaga aaatacatta 44280
ggctttccat ttcttctaat tctggggcat ctgatgaagc tttggatccc ccaaggtaag 44340
agctggactc tgctggtgaa aactctttag gaaaaacaaa agaatattgt cagaatctga 44400
tgcaccttag aaatgatgca gcagaactgc tttattttct aaaaggtgaa atggagaccc 44460
agagaagcaa agtgatttgt tcatgatcat acagctattc agtaaagcca ggacttctgt 44520
gatccactgt cctttcctta aaccagtggt tctcaacctt gggagcttta aaaaactgct 44580
agtgttggat ccatctcaga ctaattaaat cagaacccat ggggatgagg cccagacatg 44640
agtgggtttt ttgttctttt ttaaaaaaaa gctccctagg agatttctca aagaactgaa 44700
aatagaacta ccatatgatc cagcaatccc acttttgggt atctacccaa aggaagataa 44760
attattatat aaaaaagata cctgcactca aatatttatt gcaacactat ccacagtagc 44820
aaaaatatgg aatcaaccta actgtccatc catggatgac tggataaaga aaatgtgtat 44880
atatacacac acaatggaat actattcatt cgtaaaaaag aacaaagtct gtcttttgca 44940
gcaatatgga aggaactgga agccattctc ttaagtgaag caactcagaa acagaaaggc 45000
aaattccaca tgttctcact tacaattggg agctaaataa tgcatatgca tgggcacaga 45060
gtgtggaata atagacattg gagactcgga agggtggggg gaatgggaga gggtcaatga 45120
tgaaaaatta cttaatgagt acaacgtaca ttatttgggt gatgaataca ctaaaagccc 45180
acactttacc actatgcaat atggccatgt aacaaaattg cccttacacc ccttaaattt 45240
atacaaataa aaataaataa ataaaagctc cttagggctg agaactactg ctcctgtcct 45300
atgggtcccc agctttattt taactcaaaa tgagtttaga aaaatttatg aacccattta 45360
aaaatattta ttgagtatct cctgtgtgca aggcactgtg ttatgttaag tggctgaagg 45420
gaaattagac tggggaaaaa gacaaggtca tggcctaggt ttcaaactaa tataaaagac 45480
ataacaaata agaaaggatg ccaccttctt ccaaccctca tccctcttcc ttttgacagt 45540
tgcagatgtt gctaattcat tttggcaccc tttttctctg acccaaatat agtcttataa 45600
acctttttcaa cccacggctc taggcaagta tcacctttg ctcttttggc accagatctc 45660
ttgaacacta tttactggtt ttggaaagat tatacatgta tgtctggagt tgaatgactg 45720
aacagagcaa taataagagt taaagcaaga aagacaggcc tacaggagat ggcagagggt 45780
cttgcctgtc aggcattgat tttgaacttc attgcatagg caatcaagaa ctattgaagt 45840
ttttgcacaa aagactatag atgagattaa cctggttacc gtaaaggaca aagtgattgc 45900

aggtagaatg aggccagctt cataaatgaa tcatcaggat atgagaagca agggcttgaa 45960
catgagaggc catagtggga atggagggaa agggacaatg tgagaagcag tgaaggagaa 46020
gggctgattg agtaaagcag tggagaagac agtgaaagat gtcagatgac taccatgttt 46080
ggcgactgag tgagggaaga ggtggtgatg atattactga agagagaggc aaggggtggt 46140
cactggattt agagcagaca ttatcaactt gtggtgtcca gacatttcac cctgggagaa 46200
acctgttctg aagtggcttc agcatctctg aggtcagatt cctagttcta ctatttttct 46260
actgactgaa atggaaatcg agtaggcaag gcttttgatt tgtctcagtg gtctcttctg 46320
taaaatgggg gtgtttatat ccatagtctt atcacagggc tatttggggg attaagtaag 46380
acaagtgtgg cagagctttg taaactgtaa tacactgtgt acaattggat aattatggat 46440
tcttctgact catccacatg gatgtctgct gaccctgggg gaccggagcc tgggagggag 46500
gccagacctg gaaatggaaa cttgaaaatg ttctctgtag aaaagataat taacatttga 46560
ggatggttaa gtcctcttaa atagatgtca gaaaaaatgg aggtcatgta gacagaatgt 46620
tggataacac tactttgtaa aatattttat cttatttcca ttataaaaga aaaaaagctg 46680
ggctgggcac ggtggctcac gcctgtaatc ccagcacttt gggagactga ggcgggtgga 46740
ttacctgagg tcgggagttc aagaccagcc tggccaacgt ggtgaaaccc tgtctctact 46800
gaaaatagaa aaattagccg ggtgtggtga caggtgcctg taatcctagc tactcgggag 46860
gctgaggccg gagaattgct tgaacccagg aggtggaggt tgcagtgagc caagattgca 46920
ccattgcact ccagcctggg cgacaagagt gaaactccat ctcagaaaaa aaaaaaaaat 46980
agacaggaaa ataaaaaaag ccacctcaca tagtctacta ccaccaaaca catcattaac 47040
attatatttc tttattccat gctctttgtt tttaatataa acaattactt ttaagggaaa 47100
atgagaaaag gagagagtga taagacttta ttttaaaagg tggaataatt ctaaccatgg 47160
agagtattta taaatttttt ttttttgaga cagagtctcg ctctgtcacc cagggtggag 47220
tgcaatggcg tgatctcagc tcactgcaac ctccacctcc cgggttcaag caattctcct 47280
gcctcagcct cctgagtagc tgggattaca ggcaaccgcc accatgccct gccaattttt 47340
ttttttttt ttttttgga gatggagtct tgctgtgtcg ccccaggctg gagtgcagtg 47400
gcatgatctt ggctcactgc aagctccgcc tcctgggttc acgccattct cctgcctcag 47460
cctcccaagt agctgggact acaggcgccc gccaccgcac ccagctaatt tttgtatttt 47520
tagtagagac agggtttcat tatgttggcc aggctggtct tcaactcctg gcctcaagca 47580
atcctcctgc ctcagcctcc caaagtgctg gaattacagg tgtgagccac cgtgccaggc 47640
ccataaaata ttttttatag acaagtgaga gcagaaatca caggttctta tgagcaggaa 47700
aattttgaag gtcatctact ctgaacgttt ttttgtttgt ttgtttgttg ttgttgtttg 47760
tttgtttttg cttagtttac atttattaaa tacccgttat ggtccaggcc cttggctaag 47820
cgccatccat gcaatatatc acaagatatg cccagcaatc ctaggaggta gggtttatta 47880
ctacccatcg tacagaggag gaaactgagt catagagttt tagtgtcctg atcctggtca 47940
cagagccagg aagtggcaga gcaggccagg ccaagtctgt ctgacatcag agctcatcag 48000
agccctcccc attgtccttg aaccagtaaa gatggagttc ttctacaggg gtggttgggg 48060
gacaaggacc ccatgggtgt gtctgagtca gaaacatctg cgagtgggct gagaaatgag 48120
tcttctgtga aaaagagcaa aagaaaaaat gggtcaggag ccaataatca ttgtccatct 48180
ttgtgtgaat gtatggtgtg ggagtgggag caataaacga ttctaaggtc acacagaaaa 48240
gatgccacct tctccaatca cataccgccc ctcgtccccc agttttctct gaaatagctc 48300
ttcttttggc tctatcctgg cttcttcaca caggggtgtc cagtcatctc atcctggtgg 48360
gacagggata gagctgtggc agtggagatg aggaagctcg cctcctaagt gagtctgaat 48420
tcttaaatat ggagccactc cataatcatt tggagtgaat attgggccat ggccctttt 48480
cttgccagct gagctatgaa aaaaggatgt cctaagacca gaggctgtgg gaccattccc 48540
agcccctgca ggaatcaaag gagctgacag aattgtttgt ttgttttttt cacaaattga 48600
aaaaaaaaat gtaaaatttt tgaaaagaaa gcctcattga aaagaaatcc ctctccccag 48660
ctgggctccc aggcagcctc ctgcagaaca tccttagcat tgcagagttg ttcccatggc 48720
aaccgagtaa ggggcttttt gttttcctta gaagattgaa tcctttcaac cagaaggtaa 48780
ccactggttc ttccccacaa tccacactcc aaacccccta cccttatttg actacatgac 48840
tagttttgca tttatggatt tttttatgcc taattgaaaa aggctaaata tacagaaact 48900
gaggctgaag tggtttaagg aggcaactgg cccagtggtt tctcagcaac cacatgtcaa 48960
agctgtggac gttagacttg acgagagcaa gacatatcag aatctgtagc aggagcatct 49020
agtctcccag ttcaatagtg tccacaaaag aaatccagag gtttttgaag caaggaattt 49080
gggtggcact gctgtgagaa acaatcacct ggctcctcca tggggcatag agtggagatg 49140
cttcttcaaa tacccccttcc tttccaaggc catgactcag aatgactggc gtagggagcc 49200
tggacctgat ctcttcaagg aaggggaatc agatgagctg tttaatctct cttgtaaaat 49260
gaggggttat gagaccatag gctcattttg gggggggtct aaaatgcagt atttttgaa 49320

ctgatatggg gaaaaaaaga catttctgaa ttgttgtcat gttgcagatt ctgggccgtt 49380
ccagcataag caccttttctt agagtacttg gctttgtgaa gtagtcctta tcccctcctt 49440
ccactatttt acatcaagtt aaaatagagg aagatgccta gaaatggccg tatagacaga 49500
gaaaactgca ctaaaactcc ctccgtcatg cctgactcct ctctagacta tgaccatcga 49560
ggggccagaa atcatatctt aaagatcact gtgcctccag tacccagcac ggtgtttaat 49620
aaatgtttgt tgaatgaacg aactagtaaa attttcaaat cattagagct gaagtatcct 49680
ttaagattct ttagtccctc attttacaga taaggaagct aaggctcaag acattgtgtg 49740
gcttggccaa aggcacacag caagctaaag gcagagggag gacaggaccc ggctgtctca 49800
acccccctggc tgctacactt cctgcagcat ttctaattct tttaccattc ttgcgaggga 49860
ttttacaggc atgtactgct agagccgaaa taattagaag cctcttacta ctcatcagaa 49920
aagctatgtg agcccctagg gaggacacag ctagcctaga ctctgcctct ttgccctctg 49980
ctgcttatta gcagaatgta agtggttgtg tatgatgatt agtgtaagta ggatgggcaa 50040
atgcacacct ttcccacctt caaactcaga agttgtaacc aagagtcaca ctgactaaac 50100
actccaattt ccctttctgt ttttcttaac atatgtccta ttttaccaat aatagccatg 50160
gtatattagt catggtattt cacgctagct gcagaaataa cttccaaatc tcattggctt 50220
actcagtgaa agtttatttc ttactcatat aaagttgaat gtcctggtca ggcagttatc 50280
taagccacaa cttggggatg gggatgcagg cagcttccat cgtattggct ccaccattca 50340
gggatggcag agttgctctg gcataatcca accaatagag gggggaggtt tggcacttgt 50400
cagttaacca cctagcctag cattgacaca caccacttct acatacactc ccctagtcat 50460
cattcagtca tgtggcccaa cctagatgca aaggcatctg ggaaatgtag cccctatctg 50520
gtcagcaaca actttgcact tggaagggga gcctgaatcg ttattggtct ccaacacatg 50580
taactagcaa ttatacagaa cgttatttgt caggcaatgt gccaagaatt atttcattta 50640
atcttcacaa caatcctatg aggttattgt cctctttaac gtatagatga aaaagttgat 50700
ggtagagata taacttaact aatgcaaagt tgcataagtg gttggtagca aatccaaaat 50760
tcaggctgtt ctctccagag ctcaggctca tgattgctgc attctactgc tttgagcttc 50820
tgatctgaga aaatgcatca gccactaagt agcctgtgta gtctccagca attactttcc 50880
tccctctgga tcttggtttc attctctgca aagtgaggat gtttaactgg ataaaatctg 50940
atgtcacctg ccagctggga catcatatga ttctcagggt aagcatatca ggtgggtggg 51000
gtccccagtg atgcttgacc atagcaaagc cctttcaaag gtttcttagc acaccacata 51060
aatggaagcc tcacagtgtc catgtaggag aaagcagggc aaagtatttc catttaccca 51120
acaaagaaat caacatatag taaaaagaga gtgttttccc accaaggcct cagattgact 51180
agcggtagcc ttggaaatag gactttattt tgtatagtac ttttgccacc agggtggggg 51240
ggaaaagagt gcttctttgc cccaaatgct ggtttcataa aacctaaaga tgtcacatgg 51300
aaacacacca ttcccccaat ccccctcaaa aaactacttg cacttaaatg aaagagtaaa 51360
gctgtaggac tttactgagc agtgtcctgt ggggtccttg cactgccatg ctcttgaggg 51420
gctcgaggtg tatgaattcc ccagcattac ttctccttag aggtttcaga tgagcagtat 51480
gagctccaaa ctcatgctag acccaagtat ttcatgaaag aacaatcctt gaatgacttt 51540
atacagcaaa gctatatttc actgtgtcct agaaaaccaa ttgtgtgtgt ttgtgtgtgt 51600
gtgtacaact gcttgtgttc tttctaccta tgtccccctg atgcctccac acagaacatc 51660
ccaaactcca tttcaggttc ctcttgagat tcccaaactt ggaaacagga gatgcttcaa 51720
aggcctcttg gaatgtcttt tgaggcttta tattgtgata tgttggacag atggttaaga 51780
aacagaagaa gagcatcacc aaaaggattt ctcattttat gtggagatct attaatattt 51840
gccactagca aaggcattct ttcttgggaa tgaattatgc ccctagaatc agattgaccc 51900
cacagaaaca agggagaata aatagagact tgagcttaga ccttacaaca tggccagagc 51960
tgaaaaggct gagctctagg cagagaagat gcaagagcag cttcagaaga cctgagagct 52020
tatttgggta ggttcctctg gtgtaaaggg ttcttgttca cgttttcttc cagaataaga 52080
aaagaacgca aggtgtcaga gggtggatgg aaacagggta taaagcagga gcatttggaa 52140
tctgcccttt gtagcctggc ccagagagcg tcaggcagct tgttgggtaa taagtaacac 52200
tggcattttt cccatggttc tgtcatctta aagagcagga tacataaagg gattcagatg 52260
tcttgttggt ttggagaagc ttctttttaa taccttgttt taaaatttac ctggaattta 52320
ttttaatcag gtgtggtaag atgcacagac atggagatga cagtcatgaa ggaagaagta 52380
tttatactca cagatccctg taaataggaa gcatggcctc catgcaggcc aatggggaag 52440
caccagggtc agccgcaagg cagaaggagc aagaggaaaa catggacaag aggctctact 52500
gtggattcag tggcaaagaa tgggaggggc agagtaagca ggtttaggat tatcgggttt 52560
gaatgacttg attgagctgt agggtgtaga gactgcctct actgtctggc accaggggta 52620
attagggcag ctggatagtg gtctggagtg tgagagctcc ctaaaggagg tggttggagg 52680
tgtaggtttt ggattggttg atctgtatat gaaaggtgca cgtgcaggtt gagtcctcta 52740

ctatcactag aaattggctg gtcccaggag aagtagtctc tctagagaca gcaatgcccc 52800
agatgtcaaa gcatcagaaa atacagaaaa aaaattaaaa gcatgattaa ttcatactca 52860
caggtctagt ttttgtgtag ttaagagcaa cctaaagaag ttgataactc gtgttgcagg 52920
tcaggtttcc cagaaatcat attctcagat gaagatttgc atgaaggagg tttaatgctc 52980
aaactaagcc ctaaggctcc atacctgtgg aggaagtgaa agaagcccaa ctgggcacag 53040
aaggtggaac acaatgccag tcacacaaag acctcagtgg atcctgggcc atgaggagct 53100
ctaagcacag atgacccttc agaaatgtct ccaagtgggg aaaggaatca tgctagtcac 53160
tggatgtggg cttcccactc caccccatga gggcatgacc ttaagtgaga gagctctttg 53220
gacacagggc atcctaagag gggcactcag cagccacatt gggcaccaag actctcagca 53280
gctagaagaa gaaggtatag tcccaaaggg gaatctgggc tgcacacctt agtatccatt 53340
agaactggaa gtaggctgaa tcccaggcag ggatcccctg gagaacacag gtaatttttt 53400
aaaaaatcaa gctatgtgtc tgaggctatg tggtaagaca tctcagtttt ctgctaggaa 53460
aagccaccaa accagattgg cttattcatg ttgaaaagtc tgagaatcac actcagatgt 53520
tgttgataat tctgcttgga taaaatttat ctattggtat gcttgtgata tagcagtacc 53580
attgctaaaa attccatgcg gagaatccaa tctgcatcat tttctttctc aatgatttgt 53640
ttttaaaggc agaggttcgg ctgtgcccct ttaaaccttc tgtgcaagtg ccagcttcct 53700
ttcaaatgga gaagcagcag ccctgtcaga aagggtggct ggagctcccc ttttgtgaga 53760
ggaggaaaac ttactgggaa ttacctgttc gagagccaca catgaaggca taccactgct 53820
tcctctgacc ttccagccgg tatattaatg acatactgtt gtacctgaga accaatgatg 53880
aagtgggtga tgtgcctggc accttaaagg cctgggcctg ctttgacagg ggagatgata 53940
cacaacatgg ctgttagcca gctctcactg catctggaag caccatgttc cttagagcca 54000
aagttctcaa actgtgcttc ctgctgggct ccacagatcc ttcccgttcc accctgcaca 54060
caaacgtgca cacacataca cacacacaca cacacacaca cacacacagt gttctcaatg 54120
ctcgccattt agttagtatg caccaaatat gtgtagtatc tggttccacc cctggcctct 54180
cagacaatta ttagtatttt tgggagcggg gaggagagtc aggaagaccc aagcgccata 54240
tttattattt ccccagccac cccggcccag gctacatcca agttcaaagt ctatgacccc 54300
ctctctgagc tttcagcact acctcccttt gtgggggagg ggggtgccaa ttctctttct 54360
tctcatcatc tcctgttgca aaataaaagc ctaggcattc ctttgagaaa cttgggcctg 54420
gcattggaag gcgtctgaca aaggctttgt taaatgagtg gagggaggga cggtctggga 54480
gatacttttt caggtggcat aggacctccg cttcttccct tctcacatga gaaggaagat 54540
ttttctagaa atctacaggt gtttaagctg gaatgtgcct cagacatcat ctggttggac 54600
cctttcattt tgcagatctg aggcctagaa agatttggta acttgcccca ggtcacagtt 54660
gacagaattg ctcagtgaaa agtccagcat aaatacccca gcccatgtgg ccactggctg 54720
tgtgctcagc tagtgaggca cacttacttc ttaatttgtg ccacccactt ttcaggctcc 54780
cttaggacag cctccacctg ctcctactgt gcttcccatc gtccctctcc tcaggcacag 54840
gctgaggagt aataagagca cctgatatgt gtcaggcctt actgtgtgct aggaattgtg 54900
ctaagtactt cctatgaatt ttccatttat tctttataat aactttgtaa agttagagcc 54960
attattccag aagggaaaac cgaggcaatg ggagtcaaag caaagaattt gggcttttaa 55020
ccattacact attttgcaca agtagccagt aatgaaaagg ctgctatccg gaatcatctt 55080
tgcaaaaggt aatttcttta gcactttatc agaagaaggg ggctccttcc tcaaattctg 55140
agggaagaga agtggggaag aaaagatgac tgaatccaaa gctcgggcag ggaaagcaca 55200
tcgagtgcca agtgcgctgc gctggggtct agtcctgact cagccgccat cttcccaagt 55260
gcttcctgga attctctcct ctcgtggggc ctcagctcct tcatcttagg aaagaagggt 55320
aaagatctac agacaaattg atctttaagt atccttagag cactaccatt ttcagaatct 55380
aggattctat atccttccaa ttatctctgt gtagggaatt attggtcgtg tctcctgatt 55440
agggagccgg acactcgtct gtcagcccca cctggctctg caaagtccct tgtgtatctg 55500
ccctgcctgg tcacgggaga ggaagagaca aggaaacacc accgctccga ctctgtggag 55560
cacgcgctct ctcccaccca cacacccgct caggagagga ggaacctgca catttgagtc 55620
tcctcagagc ctctgcagac tcccagcagg ggtctggctt tcctctcagg tagcacagtc 55680
atgctgtaaa ctcatttggg tcttgcttgg tatgataatg cgtttagttg aagggttata 55740
taattgcaga gtcgatgatg atctctaggc caatttaaag tcaaagctat ttttaatgga 55800
attgccagag gagggcaggg atgggggcag ggaggagaga tggttagaga gtgcttttga 55860
aaccaacctc caacaatttc agccattgca tttccgaacc tgaattttca gggcagaaat 55920
tggacaatgc caattaaatc agagcaggtg tatgtgagag ctgggttcac cttcttgcag 55980
ctacagtttt attttgaata ctgttgcagg tagtgaaaat atgactaggc tgaataagag 56040
atctcagtct attcccagct cagccaaaag cccttagtgt gtccttgatc aagttacttc 56100
ccctatccat ttccttacct gcaaatgaga agcttgaacc aaactatcct aatgtccctt 56160

tcaactctaa aatcctagat gatcctcaga tgtcaacagt gctgaagccc agcactgtaa 56220
gatgtcaggt ggtccgcaga gggtgaggct cttcctgctc aaattatttc ttccacccaa 56280
gactcctcag ttacctctgt acacaacctt gcaggcccat ctaagtatcc aataacctgg 56340
ggctttagtt tacaaatttt cttggggaag aaggtaaaag ggatctagct ttctgggtta 56400
tgaatgccat gtagggaggg catggtttga gttagtcctg gtgctgggag ttcatgagac 56460
ttattctcaa atcttcagag aagaaaattc cgtgaacacc tgggaacatc aggaaaaaaa 56520
aaatgtcccc taggctactg tcaggttagg ctgctggttc tgatttgacc ttgaacttgc 56580
tataattgaa caagataagc atgtgaccta atgaaatact ttaaaacttg tagcttcctt 56640
cagcacagaa gtggctctct gaaccaattt taagcaatcc tggctctatc tgtgcatgtt 56700
gatttagcct gtggttatag tgttaacaat ttagtgattc acctcatttt taatctctct 56760
ttccctttag caggatcatt ttctctgtgt taagggatca acattgaggt aagaatggct 56820
aaataatagc atcttctgga atacaaatga ctttataaat aaaagaagat aaaaggaaga 56880
agtaggatga tttctcagct ctaatacact tagcaaatgc catatgcttt ctcctgcgtg 56940
tactggtcag gccagttcta gatacaatca tgcgctgcat aatgatgttt tggtcaacag 57000
tggattgcat atgtgacggt agtcctttaa gattataata ccatattttt gctgtgcctt 57060
ttctaggtct agatatgttt agatacacac atacttacca ttgtgttcca attgcctaca 57120
gtttccagta cagtaacctg ttgtacaggt ttgtaaccta ggagcaatag gctataccat 57180
acagcctagg tgtgtagtag gctataccac ttaggtctgg gtaagtacac tctatgatgt 57240
tttcacagtg atgaaacttc ctaatgacaa atttctcaga atgtatccca gttgttaagt 57300
gaggcatgac agtactatat ctcaagactg tccccaagct gaagtctcca gtggacacaa 57360
agaccaatgt atttagttga atcgtggacc ccaaaagttc aagtccaccc agaacctcag 57420
aatacaagtt caagtccacc cagaacctca gaatacaatt ttatttagaa atagggtctt 57480
tgcaaatgta gtaagttaag atgaggtcat accagagtaa agtgggccct aaatccaata 57540
tgactagcat ccttgtaaga aaaggaaaag gaacacagac aggggagaag gccatgtgag 57600
aacagagaca aagactggag tgaggcatct acaagacagg gaacaccaag gattgccagg 57660
agccaccaga agctaggaag aagcaaggaa gcatcctctt ctggggcctt cagagacagg 57720
atggccctgc tgacaccatt gtttcaaatg tttagccttc agaactgtga gacaataaat 57780
gtatattgtt tcaaaccatc cagttggtgg tactttgtta taggaaacta atacattcag 57840
gatggagagg tgtctgggaa gcccatgaga acaaatggaa agagccagaa gccctcaacc 57900
ttggctcgtc tacagcccat tttcttcatt cccgcatcca ggctttgaga tgacaggaag 57960
ctgtgaaacc tgtgaattgt ctccaccgca aatcctgctc cctggtccca cctagactgt 58020
cagggttgtg tggcaaggct ttcatgcctc tcactgactg cctagtacgt cccctcaatg 58080
actggtccac atctttctca cctttctcat gcatggcccc agatccaccc cagtgcctcg 58140
tcctcaagag gtgatttatt ccgagacact gatgagagca ctgtccttcc tgtgtctgag 58200
ggaaggcatg taactcttgc ttatcttcac ctgtgctcta gatcctgacc ttctctggca 58260
acctcaggga ccttgcacca tccattcttc tcgcctaatg gcgagactca gtctctccct 58320
ctccctttcc actctccctt gccattctta gtatctttct acaagcaggt cttccaaagt 58380
actgcttgag gtctgagttg gagggaacat gcctctaccc tactaaaaag agaaattcct 58440
ctgcagaaga cccaagctga ctgacaaatc cctttactgc aactgcagct ctagctccca 58500
ccattttcct gtacttactc tcctgctcag gttccctggc attgctgatg tctttcagcc 58560
tttgtgccct ggccccttc ctcctctccc ctcatctagc actacctgtc aaaatcaggg 58620
acttacttta aaatttatcc caaattatca ttgccatcat ctccactgtc accttatcat 58680
atgtttgaat agcgtttcca tttcccaaat gttttcgcat gcactttctc aattgagcct 58740
tacgaatcct agagctgaga agggtaacaa tttatgagtc ctttgacaaa tgtggaaact 58800
gacatcacag aaagtaagtt gccagccgat atgtcactgt cttcaaactc ttctttgtat 58860
ttttattatc tcccattata ttctgcctct tgtaatgatt atttctacat tggtcatatc 58920
tttccttctg tactgatctt cgcttatgat aacaaataat aatagtttac ctttgcatca 58980
cacttgatgg tttacaaaat gcttcaaatt caacatggcc cttgatcctg aagatattta 59040
tcacttaaga atcattatcg ccattttaaa atacaaattt attacttggg ctaaattttc 59100
ttattatagt tgggataggc cttcatccat agggtgagtg cagtatttgt ggactgtcat 59160
ggcagcttaa acatttagta cttgaaaatc tgatgcattg atcatcagag aaatgcaaat 59220
caaaactaca atgagatatt atttcacccc agttaaaatg gcttttagcc aaaagacagg 59280
caataatgaa tgctgacgag ggtgtgaaga aaacggagct ttcatacact gttggtgagg 59340
atgtaaatta gtacaaccac cagggaaaac agtttggagg ttcctcaaaa aactaaaaat 59400
tgagctaccg tgtgatccac caatcccact gctgggtatg tacccaaaag agaggaaatc 59460
agtatatgaa agaggtatct gcagccgggc gcggtggctc acgcctgtaa tcccagcact 59520
ttgggaggcc gaggcaggca gatcatgagg tcaggagatc gagaccatct tggctaacac 59580

ggtaaaaccc cgtctctact aaaaatacaa aaaattagcc aggcgcggtg gcgggcacct 59640
gtatttccag ctactcggaa ggctgaggca ggagaatggc atgaacctgg gaggcgtaac 59700
tttcagtgag ccgagatagc accactgcag tctggcctgg gcgaaagagc gagactctgt 59760
ctcaaaaaaa aaaaaaaaaa aaagaaagag gtatctgcac tctcatgttt gcagcagcac 59820
tgttcacaat agctaagatt tggaagcaac ctaagtgccc atcaacagat gaatggataa 59880
agaaaatgtg gtacatatat acaatggagt actattcaat aaaaaaaaag aatgagatcc 59940
agtcattagc aacaacatgg atggaactgg agatcattgt gttaagtgaa ataagccagg 60000
cacagaaaga aaaacatctt atgttcttac ttatttgtgg gatctaaaaa gcaaaacagt 60060
tgaacctatg gacatagaga gtagaaggat ggttaccaga ggctgggaag ggtggtgggg 60120
ggcttagggg gagggtggga tggttaactg gtacaaaaac agaaagaatg aataaggcct 60180
actatttgat agcacatcag ggtgactata gtaaataata acgtagctgt acatttttaa 60240
aaaacttgag tataactaaa ttgtttgcaa ctcaatggac aaatgcttga ggggatgaat 60300
atgccattat tcatgatgtg cttatttcac attgcatgcc tctgtcaaaa catcatatgt 60360
acccaataaa tatatacaac tactacatac ccacaaaaat taaaagtaaa aaaaaaaatt 60420
aagaaaataa aagaacaaaa gtagatgtat tctacatgtc tccatattgt aaaactagaa 60480
ccagtcagtt aactttagag gaaggggatt gtggacttga tataaagaca actttataat 60540
atgcagagca gcctaatcct acaattgtca aaaagtatag tggattcttt atttatttgt 60600
ccatgatatt atagaggtca tttctgcttt aacaagtagg tgggagatag ctaggtagga 60660
tatattttgt tcttattttt tattttaaaa tattgggctg tggctggaca tggtggctga 60720
aacctgtaat ctcagcactt tgggaggctg aggcaggcag atcacctcag gttaggactt 60780
ttcgagacca gcttggccaa tatggtgaaa ccccatccct accaaaaata caaaaattag 60840
ccagttgtgg tggcatgcac tgtagtctca gctccttggg aggctgaggc aggagaattg 60900
cttgaacata ggaggtggag gttgcagtga actgagatta cgccactgca ctccagactg 60960
ggaaacagag tgagactctg ttttatatat atatatatat acacacacgt acatatacat 61020
gtatatatat acacattatt attgaaagca gccaaagaaa aataacacat tatatataga 61080
gaaagagcaa atgatgagtg actttatatg tatatatatg tgtgtgtgta tatatataat 61140
gtgtatatat atacatatat atatataggt taagaacctt cagcacatgt atacctatgt 61200
aacaaacctg catgttcagc acatgtatcc cagaacttaa agtgaaaaaa aaaaaaaaga 61260
accttctgca tgccagtaac tgtgctaagt gattaggatg caatggtaat aaaaacaaag 61320
tccctctcct taaagaattt tctatttaga agggaaaact ggtaaataaa aaataaatat 61380
ataaattaca atttgtgaaa agtgctacac atgaaagagt gctgagacag acatcaatgg 61440
ataaacttta gattgagaag ggctctgaca aagcaacatt taaggtgcaa cctgagagaa 61500
tagaagttaa acaggcagat attggtgaaa gagcagtcta ggcagaggga acatcatttg 61560
caaaggccca gggtaaagaa gatcctggta aggaaatgac agtggaagaa ggttagtgta 61620
gcaggactgt ggctagggcg gagaggcagg gaagtagttt agaatttcaa tgcaatagga 61680
aatatggaag attgaaggca gttttgcatt ataaaataat atgattgcta ttttaaagct 61740
actttatcta aggatggaag attcttaaat aaacttgtgt atacttggac cacaccacca 61800
tgagcagcag ctgctctaat tcagagcagt cctcctgcca aacgctgtgt gagacaaagc 61860
tctgattcat aaaggggcat ttttctctgg gagaaaacca gtgatccatc tgtagaagta 61920
cctgagtcta aggggagacg aagcagcaaa agaaattggc ttgtgaggac agggacattg 61980
taagaatgaa aagaggaagg gaggtgctga gccctttttc ttttttcttt ttcatttttc 62040
ttttttttt ttttgagac ggagtcttgc tttgtcgccc aggctggagt gcagtggcgt 62100
aatctcagct cagtgcaacc tccggctccc gggttaaagc gattctcctg cctcagcctc 62160
ccaagtagct gggactacag gcccttttttc ttaatccaca accttcagtt ggattttgca 62220
aatgagtctg tcttcactgt ttccattcag tggctggaga caacttggaa gagaatctca 62280
gaaataactc tggctgctca cccagttgtt tgtaaatttt tattgagact ctactgtgtg 62340
ccaggctgta ccaggcactc agatatgaca gtgaatgaga taggcaacat ctttgccatt 62400
ggagagccta cactgaagtg gacatgaggg agttgaaagc aactcttata ggaaatcatg 62460
gtaaagacgt ccaagagaag aaagatgaag ggcaaacaca tgcacggatg ccaaacatct 62520
atcagagaga aaggaatttt cagacctgac ctgaatgatg aaaggaggtt tttggaaagg 62580
aaaatagaag ggaaggacaa gggaaattat ctgggcagca atatttatct gctgtggtgc 62640
ttcactctct ctctaatcct tttccacccc agccccaaat ttgaaaggat tgcagggagc 62700
tcctgctgga gtcatttctg gtattaaaaa tgtacagaaa ggaaagcttt ggttctgagt 62760
ttgcaggctt ccctgtcttt cattcctatt gtagaaagca gcttatataa aaagatgtgc 62820
tgtgtggccc tttgagctgc tgtgattgtg ttaggacccc actggatggt attcgcatga 62880
attaatctac tgtagcatct ctacaaatca agaggctggc ttctgtttga aatgtcccaa 62940
ggctttgtgc acagggcaag ctaaatgtct ccctacagtg agactgaaaa tgccttgggt 63000

gcccttgtcg ataggatctg atatatagat gcatgtctac aattgcacag tggctgctgg 63060
caacatttat tacaatctga atgtgaaatg gctattctgt tcaaggattc tgataaaaag 63120
tatcagccac agtagatgta taaggagcct ggtttcactg caactgacta cagttatctg 63180
attttttttt tctagttcat ttttagtctg tggagcaaac agagatttcc tccccaaatg 63240
atgtcctttc tcagtcacca gggtgtggtt atttggtttt atgtagagga gatagaaacc 63300
aatcagtcta aatcatattc tgttgaaatc agaaccaaag gatccacaat ctggctccaa 63360
tctaactttc cagcctcaac tcctacctgt tctttgttac tcttacccct ctaaaccact 63420
tgtgggatcc tgaacttgta acctgtgctc agactggtgc ttttgcactt ctctgatggg 63480
aaagatttct ctcatctttt atgattcagc tgaagtttca atgcttctga aatttttcc 63540
tgctcctgct ggagagcttg tttcttctgg attcccatag gtcaggtcct gtgtttggca 63600
ttgggataca aagccaagta acatagcatc catattctca aatcctcaca atttggtagg 63660
aatatagaca agtaaataca ccctgtgcaa ccttttgtaa cagaggtata aaagggtatg 63720
aaataaagaa tttaatcaaa tcaaattgaa tatgggcttc aactctgaga tcttcttcca 63780
tgatgaggtt cccagtttac tctagtgagg tcatgattcc atactggcac tcttctaggc 63840
acataaggct ctatcctatt attaaataaa gattattacc attctcactg caagcagcag 63900
caacctgaca ccatcatcat cataaaataa gtaaaacnag agttaattaa gtgtgaactt 63960
tctaaaccaa cattgtatga gataattact cataaaaatg attcttcact ttccaaaggt 64020
gcctctaaat actaagattt cagttacaat aaaacttaga tccaatttac agatattaaa 64080
tttggtccat tttccaagaa tattttcttt tctcataaaa taaaaaaagt atgtgagaat 64140
attagcacaa aggggttgca aaataaattt tatttatcca gatgtgagat aagaggcaca 64200
tgcgtctttt ttcttgtttt actgcactgg ttaggacctc tagtatgttg aataaaagtg 64260
gtaagaatgg acattcttgc tttgtttcca gtttgcttta atatgttttc tgtcagtttt 64320
tcatagatgc cttttatcag actgattaat tcagtctatt attatttcag tatgttattc 64380
agtttattat ttcataataa ttttttaaac catgaatgag tttgaatttt gtcattcctt 64440
tatgtatctg ttgaaatgat catatcgttt tgctttctaa agcttctaat atggtttaat 64500
cacatttatt gatttttcaa atgtgaagca aatttaaatt catggcataa atcctacttg 64560
gtcatcgatg tgttatcctt tttgtatgct tctgggttca atctgatact attttgttaa 64620
gtatttgtgg tgtcttttca tgagagatgt tggtctgcaa tttttttttc ttgtaaggtt 64680
tttgtaaggg tttaagaaag caaggtcagg taagcttcac aaagtaagtc aagaagtatt 64740
ttcaccttta tcttctgaaa gaatttatgc aacgttgaaa ttatttgttt cagagatggt 64800
caacagaata taccagagaa actatttgga cttagagctt ccttggggga aggtttttga 64860
taaataatgc aatttcttta atacatagta cttatatttt ctatcttacc ttgtgacaat 64920
tctgatgaat tgtgtttttc aagaagtttg cccatgtcat ctgagttgtt aaacttacta 64980
caacaaagtc tttgataata ttcctatatt agcctttgaa tgtctataag atctgtcctg 65040
atgttccctc tctcactttt ttaaagaagt cttgctagag gtttaccaat tttatttttgt 65100
tttattttat tttatttttt cttatttgag acagagtctc gctttgtcgc ccaggttgga 65160
gtgcagtggc tcgatctcgg ctcactgcaa gctctgcctc ccaggttcac gccattctcc 65220
tgcctcagcc tcccgagcag ctgggactac aggcaccagc caccatgccc ggctaatttt 65280
ttgtattttt agtagagacg gggtttcacc acgttagcca ggatggtctc gatctcctga 65340
ccttgtgatc cacctgcctc ggcctcccaa agtgctggga ttacaggcgt gagccaccgc 65400
gcctggccga ggtttaccaa gtttattaat cttttcaaag gactacattt tggctttgat 65460
aatttttcct attttttatc tacattatac tgattccaat tcttatcttt attcttttct 65520
tccttctctt cactttgggt ttaatttgtt catttttttt tctggcttct tgagatagaa 65580
gctgagatca ttgattttga acctttcttc ttttctaaat aagtgcattt aaacttacac 65640
atttcccttt aagcactgcc ttagctgtat ctcacaaatt ttgatattgt cttttcattg 65700
tcttttattc aatatattct aatttttctt gtgatttctt ctttggccca taggctgttt 65760
agaaatatgt agttagtttc caaatattcg aagactttca cagatacctt actattattg 65820
atttctaatt taattctgct acaatccaag tatatacatt ataaagtttc agccttttga 65880
aatgtattaa gaatattacc agagataaga agataagaat attaccagcg ataagtaggg 65940
atatttcata aataatagac gaattgattc atcaagaata tacaacaatc ataaatgtgt 66000
atgtgtctaa taacagagtc tcaaattata tgaaacaaaa ctgacagaac taaagagaga 66060
aatggccaat cccacaatct ttatctttat caggtgattt atcttggtga acattccttg 66120
tgctcttgaa aagaaagtgt attctgtagt cattgggtat aaaattctat atatgacaat 66180
gaggtgattg ataaaattat ttagattgtc tatatcctaa gttttgtaga attatttcat 66240
gaattactat gacaaggatg ttaacaacct acagctatga ttgtggaatt ggctatttct 66300
ctctttagtt ctgtcagttt tgttccatgt aatttgaaac tctgttatta aacacataca 66360
ttcatgattg ttgtatcttc ctgatgaatt ggttccgtta ttatttatgc aatgtcccta 66420

tttatctctg gtcatattct ttatcttgaa gtcttttaa ctgatatgaa tgtagccact 66480
tcatcctttt tatgcttacc atttgcatag tttatatttt tccattatct tatattcaca 66540
ctatttatcc ctttatactt aagtccatgt cttgtagaca gtatgcagtt aattgtgtct 66600
tgattatttt tactcctttc tgacaatttc tgcctttcca tataatatgc ttatcaatac 66660
agttggagtt aaatctaccg tcttgttatt tgtcacatct cccatctttt gttgttgttc 66720
ctcatttcct tgtttattac cttcttttca gttatttttt ttttgtattc cattttaatt 66780
cctcaattgg ctttatagct atatatcttt gtattatttt ttattgtttg ctctagggat 66840
agcaatatgt atacttacca cagacaattt agaaatcata ttgtaccact tcacataaaa 66900
tagaagaagc ttgcagcagt ctatgtccct ttacactccc attctttgtg ctattgtttc 66960
cgtatgtatt acatcacgta cattgtaaaa tccacaatag agtgttataa tctttttcca 67020
aatccttgtg tgaattaaaa attttatgag tagaaaaata catataacat tttattctta 67080
cctacatact taccagttct gctttctttt cattcttacc tgtttcagtc ttatctgtaa 67140
acccgttttc atttggtgtc atttccatta gcatttcagt gcagaacttc tagcaacata 67200
ttctctattt ccatgtatct taaaatatct ttattttgcc ttcgtttttg aaatatattt 67260
taattggaca tagaaatcta ggttggcagt tttctcttat actcttgggt ttcattgtct 67320
tctgatttct gttgtttatg aggaaaagtc attgattatt tgctctttct ctatacacaa 67380
tgtattattt ttctttggct gtttcaagat atttttctct ttatctgtgg ttatcaacac 67440
tttgattatg atggcctaag tggtattatt gttgtttgta tttattccac ttggtgttcc 67500
ttgagcttct aacttctgtg agctttttt ttctcagcga atttggaaaa atttaagcca 67560
attattatat aatttttctt ctccattctt tctactctct ttggaactcc agttgtacat 67620
aggttagact gcatgacgtt gtcccataga tcactaagac tctgttcatt tttcaatttt 67680
tttctctatg ttcttcagat tggacaattt atcttgatct ctattaatgt tcacttatcc 67740
tttattatgc caccttcaat ctgatattaa ggccattcag atctagaatt tctattaggt 67800
tattatttat agtattaatt tctctgctaa gatttttgt ctgttcattc attatgacca 67860
caatattagg ttcttaaaca tattttaata gctgctttca agtccttgtc agttaattcc 67920
atctgagtca tcttggggtt attttctatt gagtgatctt taccttatct gtcggtcaca 67980
ttttttttctg tttcttcaca tgtctagtaa ttatttattg tttgctgtat attgaaatga 68040
aatattataa acagtatcaa ttacattatc ttccttttaa gggtattgag ttttgttctg 68100
gaagtagtta aattactagt agaacttttt gttcctgtca aacttgatct tattctttgt 68160
tacagtgagc ctattttagt tttaaagtta gtcctagggt acaactcttg ctctattgta 68220
tgctccttac ttctatcaca tttatttcta ttgcctgaga tagtcaatga gttctcacct 68280
gagcaggaac tgcaacattt cttgacatgg tcttacctat gtattcatca ttcatctctc 68340
aggcctgtaa gaagagatct ctgttgggtc ctgtggaatc ttgcttgcac ttggacagct 68400
cagccttcag ccaaagactt gcaggaaaac cccatagaaa catctgggcc ctctcaatat 68460
ttgatgttta ggaagctaaa cgtcaagtat agcctccttt tctagggacc ctatcttgtg 68520
aatttcactc accttaacaa ctcagaactc ttatcttctg ccttctcagg ggagctaaac 68580
tgtcactttc tgtgggctcc atcttcctgc tccacaatag gaaagtatct gcagagaaaa 68640
ggctggacaa ttgtgtagta attgcttcac gcatttccct tctctcaaag attgtaagtt 68700
tgcactgttt gctgttcaat acctgaaaat gatttctaca aattgttttt ccagttttat 68760
gattgttttc aatgggagat catttctagt accagttcct ccatcatggc cagaggtaca 68820
agttcaactt ggatcatttt aaaaatacaa actggggcat gtcacttcct gccccaaacc 68880
ccttggtagc tttccattgc tcttagaata actttgtgat ctacaacatc ttcttcaagg 68940
ccccgcatga tacaaattct ggctatttct ctagtttctt attgcaccac cttgtccctc 69000
atccaccttt tttttagtct tctctctttc tttgaacttc taccaccagg ttttttcaca 69060
cgttcttctt tccccattaa caatgatcca ccattctctt tctttatcca ctgttactca 69120
tcctcataac tgaaacatca tttcctaagg atggccattc ctggttcagt cagtctatat 69180
ttcatccccc atcacatact cttgttttac cctatatttt tccttcaaag cacttattta 69240
agttgtaatt atgtgttgtt tattttatgt ctgtctgccc tcacagaatc cacagtccag 69300
gagaacagaa atcctgcctc ttttatttat accacatcca cagtattatt agtgcctgtc 69360
acctagtagg tatgcagtat gtacctattg aataaatgaa ttgacttctg tcttttagat 69420
cgtctactca ttttatcatt gatgacaaac ataatacctt acattcgtgt agtctttttc 69480
actcctcaaa gaggattttc tgcatagctc ctctgagcct cacaaaaccc tttaaggaag 69540
attgtgaata ttatcagata aagattgtga gacacagaaa agccagatga tttggcaatg 69600
ctcatagtac cagaggcaga aatacagcta gaacagtctc ctggcctcta atcaggagtt 69660
ctttccagaa cactgcttca tcttccattc tcttgggttc tttctatcct tactttatag 69720
ggcaaaatgt gtgcaaagta taatccctct tttgcaatgt gtttttagtt tttcagattg 69780
gaatcatgta ggctttttat gcccttaata aatatcagtg agcacaaagg aagtcctgtg 69840

agggcttata atcattttgc tcccattaat tccaacactg agcagtttcc ccatttccat 69900
tcttggcctt gtgaagctct ttgctatccc tgttaaaatc taaagttgct tgaaccttct 69960
tattgcaaaa atgcatctta aacattctaa tacctctttt ttaaaaaacc aataaagact 70020
acgtcaaaaa tcagccatca atcgagaagc cctgcagtca tttgtgtgct gttgtcccta 70080
agtagaagtg aatgtgctga gctctgcatt ccccacctag ctcctctgtg atcagggtgg 70140
acattcccag gacaactggg ccgaggctgg aaacaccatc tgaatgtctg accacacaaa 70200
gttgagtggc tgatccaggt ttaaccttga cctcatcagc accaccttct aagcaacact 70260
ttggctcaga agcccagtta tttattccaa gggatgattg aatgcagtgc tagtgtttct 70320
tcagggcttt tgaactcatt tatttatcca gtcatttata aaagatgaag aggagaacaa 70380
ggtaggccaa agtggctttg tactattaaa ggctgcttga tttctaagta catgttcttt 70440
gccacctttc tgccattcca cattctagaa gccatgggta agtcagcaca gggatcttaa 70500
catgataaca ttggttttag gaggtctcgt gcataatgga ccagacttag agcacaatgc 70560
tgtaaggtag tgatttaggt gagcagcaga ttctggcttt aggagtttat tatcagatgc 70620
tttttaaacg acttgtggcc caggatccct gcacccatgg gaagcattgt agccttagaa 70680
ctctgggaat tctgaatata attcctgaat caatcgtaag gatgcatatc tgatgcttag 70740
tgcaaaccaa gaggcagaat atttgcaggc agtgtatcct tgaaaaacaa atctaggtca 70800
ttttcctgcc atgcttcaag cttacttttc catccttcct gatggtagta ctaactacat 70860
ttgtagacca tttacgtggt caacactgtg ctaagctgtt agcttcattc tctatgagac 70920
aggcactctt agcccaactt tacaattggg aaaactgaga ctcaatgaga taaagtaaat 70980
tctttacagt cattatgcta gtccatgaag gagctgcgat ttgcaactaa atctatctga 71040
ttccacagtc tttgctttta accagaggtt agcaaactac ttctgtaaag ggaagacagt 71100
agttatctta atctttgtgg gcaacatagg gtctctgtaa cgtattcttc tttctgtcac 71160
aatcttctgg aatgtaaaaa acatttaaaa tttacaaacc ttacaagaac agctcatggg 71220
ctaaatcgga cctggattta gtctgtgaat catagtttgc tgaccccgct ttttaaccag 71280
tatgtaccct ccttctcggg atgtgaaaaa ttagtgcaat tgcaatggaa aatagcaaga 71340
aaatggtaag ggcctggaag aggcagcagg attacatcag gtgctatccc tgctctggtg 71400
agatgaaact ggggatcatt gaaccacctg gcatttgtta aagagttctg ctttccctct 71460
gagattcttt caggaacctc acacctctag cagcccggag aaccgtgggc tgcaaggaaa 71520
tgcctcctca aaggagtaga aaacctgcag gatagaaatc atcacatctg tctggctttt 71580
ctcaaccttt ctcttctgca ctttcttgga tataatcaaa gcactaccag gaactccaga 71640
gtcggcacct tttcattttt gtgttttcat ttaattattt ctcagctgct aagtgtttga 71700
ctgtttaagg gactctagtg gtaaatattt gtctttagcc tggcagaagc tgtggtttcc 71760
tttgatgagc tcacacggtg tggcttttaa gatgctgctg accaggacag ctgactgtcc 71820
ccagtgggtg cagtccccag cagtgggctg gaccccttcc agaaagcgct gctgggccaa 71880
gaggcttcct ccaacttccc gctgccccca tctaaccaac acctcagtct cttctccacc 71940
tgcttccctg ccctcttcct ttccctcgca gacactttct tctgcctggc aaaaggaatc 72000
ttgtttccat ggaagcctca ttaaatctgc atcttgctca gtttgggttt gatcacggct 72060
gccagaagta tttttagccc atgcagttgc gtaatgagat agagattggg gaaaggggga 72120
ggtgactgta taggcagagg gtttttttaa aaaaaagtga gaaagagaag gaaaacctct 72180
aaagaaaaga gttttatgga attggaagaa ggatggagca cctcttttgg gagcatgagg 72240
ctggtgttct ctggttagct cttcccactg gaagcccatg gacacttgcc ataatacctg 72300
tcctggtcac atgtcagggg aacctctgat ctccctttcc atgagcttag ttggcccagc 72360
cagggtgaca cttatgctag ggagtgtgat tgatgttgct gcttacagat ttcccctccc 72420
acagacctga tggggcagcc aggatagtgg cagagaagaa gacagagcaa tagcaggaaa 72480
gagaggacaa cactaacaca ttggaggttt atgttcaaag acgggatcta gggggtcaga 72540
gaaagcacac ctaccatgta attggtgctg gaatctgatg ccaagtgcac ccttggcttc 72600
tgaggttctg agaactcttg cttgtgcttt tcagccagac tatgccctca cctgcccctg 72660
tactttaaag agctctttag gctggagtgg ttgtttgcat tggattgttg gagtgtgtgt 72720
gcatgttgtt gtgttcttgt attacaagac aaagagatta aaaaaaaacc acatgcagct 72780
gtcacagcta atgtttattg aacttttact atgccacatg gtgttttaag cattctatat 72840
gtgttaactc attttcccta attctatgga ctagacactt aaacagtctc cattgtacaa 72900
acaaggaaac tgaggcacag agaggttggg aaactcattt gaggtcctcc agctaattaa 72960
tagtggagcc aggttttgta cccagacaac ctgatttgag aatctgcagt cctagattag 73020
taacgtgttg ttggcctgtc acacatttta aatgacattc tgtacacaga accatttata 73080
gtaactttgt attgttgagc tgaaagcagt ctgcagatgt gctgctggga tttcattcat 73140
cttcaaagag gtgtttttttt ttttttttaa aggaaaatgc ttttctgagg gtggtatcta 73200
aattcataaa aatctttacg atcaagattt tcacaaattt cattctgact ctgttgcatt 73260

gcccttcttc ccatattccc agttagtttg tattgattgc tgcatctccc ttgagcccat 73320
ggtcccccac aacatttctt gcagaactgt gtcctgcctt cacactgtca ggcagcagga 73380
gcctctctag cggccagccc acagtcctgc agctccttcc tcaggacgtt taatttccca 73440
catttctatg cagttacctc acagaaggat ggctacgagg gcctcacttg gcttggcaag 73500
ttggtcccct ttttactcac aagactctgt ttatctcttt gtttatcttt gtttatctct 73560
ttgttgacct gcccctcttc aaggcctcag ttttctctga agtttacagc ttccctcctc 73620
atcccgcaaa agaccaaagt ggaaaagatg aaaccagaat ccactgcaag ccccacctgc 73680
cacagcctct cctctaaatg cattctctgt tgtgtttagg acttgagaat gaagagggac 73740
atgaattgag gatttgttta ttattcttta caatatccct gtgagctgag tactgtaaat 73800
acccccattt gatacatgag taaactgagg tgtggagtga tagaggaatt tgctcaaggt 73860
cacataacta gtaagtgggt ggagctgtga tgtgaaactg ggcagtctga ttctgggacc 73920
tgtgctctta atcaccaatc tatattgcct cctacttgaa aacatccagg gaaaatgttg 73980
agatagatca gctgaaatct tcttgcacag taaagcaggg gccacctgtc ctggagttac 74040
attcatcttg ttcattgtca acgatttgtg ttcagtgaca ccctcttcag cccaagaact 74100
tacctgggtg ctgtgacaat tggacatgac taggaacaac cagtgacatt gtagcccatc 74160
caaacacagg gtaggaagtg gatgcttgtc actctctttt ggttataaga agcaggaacc 74220
cagtaaaggc acctttttata tatctataaa gttgaatata taagatatat gggggccagg 74280
cacagtggct cacacctgta atccgaacat tttgggagcc caaagcaggt ggatcacctg 74340
aggtcaggag ttcaagacca gcctgaccaa catggtgaaa ccccatcttt actaaaaata 74400
caaaaattag ctgggcgtgg tggcacacac ctgtagtccc agctacttgg gaggctgagg 74460
caggatactt gcttgaaccc gggaggtgga ggttgcagtg agcagagatt gcgccactgc 74520
actccagcct gggtgacaga gcgagattcc acctcaacga aaaaaaaaaa gaagatatat 74580
gggtatgtgt agaactcaca gaagggcaaa caggccttaa caggtgctga aaacaggaac 74640
tgggaagttg ccagtacctt cctgtctttt cccctggaac caaacggttt cttacttgct 74700
tctctctgca cctctgtctc atttccctct ctcttcagat gatttttcat tgttgcatca 74760
cacacataga aaaatcagga tccaccctcc caagtttaca tatcgttgtt tcaggcagcc 74820
atagtatcct taaaactcca cattccaggg agaaagcttg ggtcaaggat tcagccaaag 74880
ggcagcgaaa tggagtaaag atgcaactgc caggtctatg ggcagcaagg aggccgggaa 74940
ggaagccgct gttgtggtcc aagtgacaat tcaacagctc aaagcataag taagttgtgt 75000
gcttttcaca gatggagaaa ctgaggcaca gaaggaacct ggctggggtc caggtctctg 75060
gcctttgtgt caatgctagg tcactggatg tggcgtctga tttctacagg aaatgtggtt 75120
tctctacttt gtcccagagc ccactcagag cactggctgg ccaggggtc ctagggccct 75180
cttaggatag tctcaggcca acagccccag gacagaagca accaaagtga agttatgaaa 75240
gaaagctctt tgctgatctg tcaatggcac ccttgtagag ccaatactta gaacacctgg 75300
atttgaatac tcatctccaa aacctgtgtt ctttctacca cgtgacaagc ccttgtaaac 75360
ctcacaacgt ctctatgagg tgagcgcttg cagatccaca ctttagataa gcaaatggag 75420
gctcagaggg taagcagcta gttcaaggtt atgcacctga gccaggatgt ggacacagct 75480
ctgtgtctga ttcctaaggg cctgtgcttt agccactttg caatactgct gctgtctgct 75540
tcatttcctc atctgtcaga tgggaacgat aatactcaac tcacatggat actgtatgag 75600
gaaaaacaga taaaagaaga gaaagtgctt tgaaaacata agcagccctg gcagatggga 75660
attatttttg ctgctgacac acatcctcag ccttgagggc tctgctgagc catacccagc 75720
tcagagctct ggaggcacct cctccccatc aacagcaggg gggacattct gtcttcatcc 75780
tgagcaggct gacaaactga accccactcc tccctcaatg tccccatgct gggaaggagt 75840
atagctcatg ctgtgttctg tcttgttgct gagagaatgc agaacccaga atttgggtct 75900
cagcaggttg gggagaaaag gaaatgtatt tcttccccca agatttcttt ttgaaatatt 75960
ttcatttgtg gaatcagatt gtgcatgcaa gtttcttcca gaaatgtaag acgtcgtaat 76020
gatgggaact gttggtttta taattgaagg atgggaaagg aaactgatat ttatggagca 76080
cctgttctat accaggcagc tacccaacca tcagccattg ttgcaatgtt atgcaagctt 76140
tattatccac atttcacagt ctgagtctga ctcagcaatg ttgtgttcta tgtgctagtt 76200
cccacaggta ggtggctgca gcgctgggat ttgaacccat ctccaaagcc tccatctttc 76260
taccactgcc tcccattggt ggggaggcca tggactggct gtcagagatg tcctttccag 76320
tctagcagac taggaagctg ctggaagcta cttatgcaaa ggtcagcaag gaaggaaaca 76380
gagtcagaac tagatggggc tcccctggcc acttttccat gctggcccac atgtccggct 76440
agcagtcaac attgggtctt atgcagagcc acctgtgttc aatggaaaca tcctggacac 76500
tgcacaaact agtgggagcc tgtgagggaa cagcctgtcg ggttcattga ggttcagccc 76560
aactcatgag ctagggcagg taccagaggg tgtgttccac ccaaatgggg caggtaggca 76620
ggggacacag gctccatttt catgaccaaa gactgagcag agaggctctc tgagcagtgg 76680

cagaatggga agtgtcaaga agctttgttt gacaattgag tcaagaggac agaaaagaca 76740
gaaagcagac atcagagttg ggaaggctca ccccagctcc ttgacaaagg tgcatgaggc 76800
cagttcttga agcagtgacc ctgccttatg tcatgtgttt atcaaagccg gcccatcagc 76860
cctgaagtgg cctctgtgtt tagaagaggg cctgacatga ttctctgaga aaggatttga 76920
caacaacaaa gtgttgccgt atgtgttgtc tcatcccctc aatagtcctg tgaggtatgt 76980
gagacaggtg ttactctctc cacttggcaa atagggaaaa gagggcccag agaagtgaag 77040
ctgctttccc aggaccacac agctggtaaa cagtgtccat ctcagctgtt ctgtctccca 77100
caccaaatac cctgtgcacc acgcaaacac aaagacaact ggacaaccaa gtcatctaat 77160
gagtatgcat gctatggtct ctctcatttt gtctttcagg gctataccct aggagagcta 77220
atcattcttg gttagataag aaatagccaa cacttctgca gcatggtagg ccaaatacca 77280
ccagaataaa ctcagaccca aagagatgct cagaatgtgt ggagttaata cttcactata 77340
cagctctaag gtataagcct tgtccatctg tcacattatg acatgtgctt gctcccacct 77400
caattcctga ttccacatta caacaaatac aatttcaggc tttgaactaa caatgccaat 77460
gtttctgaag cccatattaa atgccaaaat ctgagtcagc tactggaggt agagacatga 77520
ataagatggt ccatattatt ttagaggatt ctttggttgc aaagggcaga cacccagctt 77580
gaattcactt tggagaaatt gggatttttt tggcttgcat aagcaaagca tgagaaagaa 77640
agttccaggg atgatgaaaa ccaggaatgc aaatgtctcc agaattcttt ctttttctct 77700
ttaggccatc ttttttctct caaactggtt ccctccactg ggctggagac gttactacca 77760
gcagcactca gacccacatc ttcagtttaa atgttggaaa tggactgtca gagaacattt 77820
aggccattca ttctgtggga gagataggct atgtaaaaag atagccactc ccatgtgaac 77880
aatgtggtta ggattagagg catgaatata ccccaaacca ggggtgtggg aaggaggttg 77940
acactctagg tgataatacc cagaccttaa ggagctttct gtctagaggg aggtatggac 78000
atggacaagt aatcaacagc tacaaagcag agctgccagc tctgcaacac aagagccctg 78060
agaggcatga caggggcagg gtggggatcc atgtgggtct ggattgaagt gaggaggggc 78120
atcaggaaag cattccagga gagctgaggg acacttgagc acaccctcaa agaatgactg 78180
ggggtcatga ggtatacaag ggaggaagtg cacccgagac agaaacaatc acataagcaa 78240
aaatgcagaa gaatatgagg atcggggaag ggcaagtagc tcagtagtgt tggaggccaa 78300
gggacacgaa ggaaggtgat aaagccctga tgttaaggat agaaaaatca aagtcctttg 78360
aaaatcatgt ggagttagga tctcaagaac cctacaagga tttctttaga atagaatcaa 78420
agaaaaacaa agtttacagt ctgtgagggt tgcataggaa gtaacgtggt gagaaatgtt 78480
ggcttgagaa ccacatatcc ataacacaat ggtgttttag aggatttggg ggaagggaga 78540
gaaaatctca aattgtctca gtaactaatg agctttcatg tacatttaaa atagtaataa 78600
atgcaattgt gaggatgatg gtgagatgag caaaataatc cagtttgtaa ttgtagttat 78660
caggctggca tatcctgcag gtcacacttc taaacatgac ttcgaaaaat caaagatcag 78720
ctaagtttga agtaagtatt gaaagaggga gattatgttg cctcaagtta aaatagaacg 78780
taaaagatgg tgattcaaat gatcaaaagc accaagcttc cctgttagga ttcaagggag 78840
gggtgcgtgg ctccgacacc agatatctgc aaagcaatat gaaatgagat caatagtaga 78900
cattgaaaga ttgaaactga tataggatat tcaagtacca gcttcaagaa aatgaaatga 78960
gacctaataa aagagagtag gagtcaaggg ggtatacgat attaaagaaa gtgaagagcc 79020
agggtttgta ggaaggaagg gagaagaggc aaagagagca gctcttttaa cacaggagct 79080
tcctcctttc ccattctccc tcctgctaaa agccgagttt gttttagctg aaatgattgt 79140
aagacaaatt tttattatta aaaaaggagc tattttgtgt tggtttccat tataaaatca 79200
gagctctgct gccataaaat taaatcccat aataaaatga gtagaaaacg tgatgtcctg 79260
cagaaaggaa gatggcagcc cactcagtgc catgctgggc ttgactatat acaagccgtg 79320
catctcctgc tgcgagttgt agctgctgcc cagcagtgca cattatcgtt gcagctgttt 79380
tcctcacatt ctgaggttta tgaaatccct catccatcaa taattgatct ttagctctta 79440
gtccaggggt tgtcaactgg cactccatgg acctttagag gattgatggc taggttttca 79500
aagatctttg aacccccctga aattatatac aaaatactgt gtgtgagtat gtgcattttt 79560
ctggtaagaa gcacctgaat tatcgaagca gtttgtgatc ccccaaaaag ctaagaacta 79620
cttcctagag caaagggaga ttttgctaca cttagagatt tacacatttg accagggcag 79680
ctcacacaag tgggatgcgg tttcacattt catggcagat ctgcttccag ctatacaaat 79740
tcatcaagga aatattgtaa tacttctata tgaatcagga attcactata tttaacttat 79800
ttggaataag aaccactata tatatacaag tttttccaaa agactgaagg ttcttcctgt 79860
ggcaggaagg aatatgatta gattcatgaa gcgcctttat gtttatattt caactctgaa 79920
agataattgt gactttacta aatcaaacct gtataccacg attaggaaaa tgtggactga 79980
tttggggttc taggggtaaa atgtgacccc tgtgaagtac caatgcaccg ttcttttatc 80040
tgtgaacggg cactgagctt ctgaaattaa ttagtaggca ggaggacatg cgcatatgac 80100

gtgatagttt aagtactgat aattattcac ttggaaggga agagaataaa attcagaaca 80160
cagtattcct taatgggaaa tcaacttaga ggaggtagga gggagatcaa gcaagaatat 80220
ttctggtaaa acatgcataa atcaatggtc agccaatgtg ttgatcaaag aaattatctt 80280
tcggggaaaa cagtagaagg caattgaaaa acaagcatca ggctgcataa aaacagcaaa 80340
caaaagtcac aatggcttga ttgtgtgatg aggtaattaa tggctgcagt tagcaaaata 80400
tgttcaaaaa aaagacagaa agggtagtta caggagaaaa acatccccgc agatcttcaa 80460
aatcagaaac aatgaaaata attatttcaa aaattaagaa aaaaactctc taatttatac 80520
ctgaattacc tggataattg gtaaaatttc ctgcatatac aaatcttggt cctctgctcc 80580
tctctctata aataaataga aatgtatgaa tcaatagtca gccaatgtgt tgatcaaaga 80640
aattatcttt tgggggaaaa ttggtagaag ccaattaaaa aacaagcatc atattgcatg 80700
aaaacagcaa acggaagtca caatggctcg acggtgtaat gaagccacac aatatgtatt 80760
aaacacatca tctacacaga tggattcaaa gataccttct ttgtgtctaa gtcccaaatc 80820
tgtgtttcct ggctctgttc cctcatatct agtcattctc caagtcagca tgcccaactt 80880
gaaagtgtca ttttcaaaac ctgcttcttc tcttctggaa gttcttcctc tgcccattgc 80940
tccacaatcc ccacctcttt cacccagtag caaaccttaa atttatcttt tactttgtct 81000
tacttcccct tcttatattc aaaatgtttc tcacttgcat ctcttttcat tcatttcata 81060
agcatttatg agctcctgtt atggtttgga aactgttctt catgctggag gtggtcttat 81120
aaacaagtaa tttcaattga gtatttagta tgttaagtgc catcccaaag gcaaacacca 81180
gctgtgggag gctccccaaa tcagtctaag gaagttggga aaagcatctc agagaagatg 81240
gtgtctgaga tggggaggat gtgtggaact gggcaaggaa gagaacaagt aacaacattc 81300
tagaaaaagg cctctttcag catgctaaga agtttggagg acagaggagt taccattcaa 81360
aatttggagg gaaggaagag catactgagg tttgccactt gaacagataa tttcagctgt 81420
gttgggtgag tgaagttgag tgggtacaaa tcaggtcagg aatataagtt aggagactgt 81480
tactagaatc caggccagag gtgatggtgg ccaatatatg agagttttag cagggaatga 81540
aaaaaagaaa atgtgttcat gaggtagaag taggtaaaaa caacaggatc tggttcctga 81600
ttggaaatgg gggtagcctg gagaggaagc cagaatgcag gcaagaatgc atagtggtac 81660
catccactga catagggatt aaaggaggag aagaagcttt ggtaaagaaa ataagaagtt 81720
cagctatgga atgtttgaat ttgatttctc tgatgaggag tagttctagg tgatgataat 81780
gctcagggtg tagacttgag agtggatggg taaagtaaag gttgaggcta ttaaaaggga 81840
aaaggtcaag gaactgaggg ccaaggattt ataataagtt atcttgggcc actaaagcca 81900
cgcaggatgc tggcaggaaa cctatgagcc aggtcttcaa tgttgagtcc agtgactcag 81960
gtgtcagaag cagcaggaga agcattgata gcctgatggg gaaggagccg ttacctgaga 82020
gtagcagaga gagttatcct agctgacaca gctctcaggg atttgcttct aaagcaatcc 82080
ttaggaaaga aagagcagta tccacaggag actggtgggc actggcttcc ccagaaaacc 82140
tacctagatg aattctattc tcaagggact cctatttaga taaggggctt tgttagttct 82200
cagagcaaca ccaaacagat gtatatctca ttacttgccc ccacaacctt tctgctctgg 82260
ccacatgggc ctacccactg tctgctaaat gcacttcata ttttcttgtt tcagtgcctc 82320
agtattcata atcttctttt cctaatctct gcccctcact tacctgaatc ttttgtattc 82380
tcaatgacct gctccatccc agccctttca agaaccttta atacctacca agtgaatact 82440
ctctccattg attacacact tcctgtagca cctgttctat aattatgaaa tattacctat 82500
tgtacacata tatttcaatc tcttggtgga cagagaatcc aatttatgcc ttgtcaattt 82560
gtagcacatt tccttgcata tgtagatgca ccatgaatat ttagagaact tgttagttaa 82620
tttcctgttt aacatgggct gcaaagttct ggtccatgca cgtcttttat aaaatagaaa 82680
tgacggatgg tgcatggagc ttaaattcca tgaagcagaa acatatgaga gatggagctg 82740
aatttgtttg cctgtacagc tcttacagca attgcttcca atttgtttga tttacctaag 82800
agctaaaatt gtaaatggca gctcaaatga tttttctgta cattcagaaa atgagtttga 82860
atatttgttg gagagtaact gcttaagaca tgaaaaaggg ggagattata gcttttaact 82920
ctttttttatg gcagagcatt aaggaaaaaa aagtgcagat aaatgagatc aaatggcaag 82980
tgtctgaacc tgctggacac aagtcccggt agccattgat agacagtgtt tatatgactt 83040
ctgggccatc aatagataga taaggtacat cagcggccaa tgttccagga agtttgagaa 83100
gataaatgga agttgcacag cagcctaaaa gcttccttag gagggctgtg ctcctccaga 83160
gcgccatctg cctgtgtctt cctgttcttc ttcttcacat taaatgcttt tccttttctc 83220
atttttatga tggttatcct aaagatatgc tagcctggac tttgacaagg acatctggag 83280
ataagaaaga ttctgaatta tttttccctt tgggcaattg tagcaatttt aaaactatgt 83340
tagatggcta gagattcttg agaatatttc ttttcttgga aaatcataag gctttggata 83400
gtggtaccta tagaagctga catcagcagc agcctgcctc cagtcgatca gggcctttgg 83460
aacttcacgg ggctcctcta ctgacagccc catcggtttc cctccagcac acgtaactca 83520

gcattgactc tgggtagtag agggtggttt atggaatctg attcatctca gaaagaggtg 83580
gatgcaaaca cattcccaga gcagaaggct tggcatgtct ggtcttaggc agagggaact 83640
ggagatactt gtcctattgt tcttgagatt ccagcaaaaa tagcccatta cagaggaaga 83700
agatatcagg tcaaatgaag gctttggtgc tacaacattg tcttagaaaa aaaaagaaag 83760
aaattggcca agtgcagtgg ctcagcactt tgggaggctg agggggggcag accacttgag 83820
atcaggagtt cgagaccagc ctggccaaca tggcgaaact ccgtctctac caaaaagtat 83880
taaaaaatag ccgagtgtgg tggcgggctc ctgtaatccc agctactcgg gaggctgagg 83940
ccggagaatc acttgaacct gggaggcgga ggttgcagtg agccaagatc gtgccattgc 84000
actccagcct gggcaacaga gtgagactcc atctcaaaaa aaaaaaaaaa gaaaaaagaa 84060
aaagaaaaaa gaaaagaaag aaattaaatt aaaaaaattg ttttttaaac aaaggaaggc 84120
tttgggcttg gagtccaact aagctaggct ggaatcccgg tttcatctcg cttctctgtg 84180
caactttgga ttttactgaa tctctcttat tctcaattcc ctcctctgta aaatgaagat 84240
aatgctagta cctgtctcat caagttgaag gagacttaaa tgagatgtgt tgaaagcatt 84300
tagcatagta tgtggcacat aaagaacact caataaatgc tggctataaa gaagccagag 84360
agagactcgg aggtgatgag agaggccaca attccctcca tttcattgaa aagcaatttt 84420
tattatctca tttgaaaggc agtatagtat agtggttaag gacatgcact atggagctag 84480
acctcctcag ttcactttct gtctctatca tttattagct gtgacttaac cttcttgtgc 84540
ctcagttttt atcatttttg agagaggagt aataatagtt cctactctgg tgtgttgtgg 84600
agatttgatg agttaataca tataaagcac acatagtagt gcctggagca tattaaatga 84660
catgtaagta ttagctgtta ttttattaaa caacatgtgg cataggacat attggaactt 84720
tgaagtcttt gaggctcttc ccagtttcat aaatcagaga ctacagtata aatatctgct 84780
tacatgtctg ctttccccat tggactgcga aatcttgaaa ctgttttatt catctctgca 84840
tagcgttggc atcgtattat gatacctgac atttaccagg tgccaaatgg gactgggcat 84900
gttgtaggga ttcagtcaat gtgggtcatt gcaggcgggg aggtgggtcg ggttaaaggt 84960
aagagaaggg ccttggggca tcacattaag tagttaccag attgaactgc aaacattgct 85020
atccaggaga aatcaggtca atatttcacc ttcatggcaa taccagtaca gtccaaggag 85080
aatgcataga aggaaagaaa tcataatctg attgtatgtg ttttttagt agtaaataat 85140
aataattatt actattccta tacaattttg tgtgttggtg tgttttgttt tgttgtgcat 85200
gaaaaatggg gtgctaatct attccccttc ccaacaccag tgctcagaag aaatttccac 85260
agatagagaa gctataggtt atgaatttgg ccttgatgga ttctgggtca ctatttctca 85320
atgtttgtcc atgtcatgtg aagctcttaa gataaagaac aatgtcttac tcgtcttttt 85380
aacttcttta ccccctaatg cctatcacat actttgccca tggaaactca atagacattt 85440
gtaaatggaa tttaatttct gaggtccagt aaagcctttt tccatccttc ccctactaca 85500
cagtttgtct aaccatgtct tcccttccat catccacctt ataaacgtta ttactcattc 85560
ttccatcaca ttcttgacac ctcccatgtc caatgtcaaa caagtaccat ttgggaaaca 85620
gaattctagg aatctggaga cctagagctc ttcagaccct gaaatccagt tttctgagct 85680
gagacagttt cttaatttct cactccaact ccgtttctcc tctttctcaa tggatatttt 85740
ccaagtctcc attaggcata tagcaattcc agaaaacatt caattttccc ttctcttaat 85800
gccatgctcc aaaacaccac attccctcta gacattgagc attggagaga gatggaaaag 85860
tactttgaaa atgtgtgcat gtgagaaaaa tgctaagtgt tctgtctggt cacttcaatg 85920
acaagtttgc tactttagaa acttgactaa acagagtgtg aggaaaaaca tgaaaagaaa 85980
aaaatgtgtt cagcttggct gaataatgac cagcagggtg aaaagataag ataaccaccc 86040
gctcacagga tttctatcct caagccctag aaggttgaca acagcagaca ctgaaactac 86100
tcttaatgga gggtgtgcta aagaagcaac attatagccg cttttaggaa agcaaatagg 86160
aaagttggtg aaatagagaa gatgcctaag catgtgagat accacctcca tcttggaaaa 86220
taaccaaggt gatacaatgt tatgcaggac cccttaatta aaacagattt agtgattaat 86280
atcaggagca ttgtcaagaa tcacaacaac agcaattagt tactattgag caatttctgc 86340
taagtaattt gcaggagggc atctcactta attatcacat ccttttatag atgagaatat 86400
agaggcttaa aaaggtgctt ttcccaatgt tattcagcta taagtggtca gtcatgactc 86460
aaacataggt caacctgaca acaagatctt cactcttaac ttctcttctg tgttgtaata 86520
cccttgatcc atggaaatgg accatcttca tatactgctt ttttgcctct ggaatgtcca 86580
ggtatggatt gggtaatgct caaagacaga gaggaataga gtattaaaaa gatccctggc 86640
ctcattttct gaagacatga gcctaagctg agctgtacca tttaccatct atgtgaactt 86700
gggcagattt tttgacactg ctgggtctca attcctgtaa ctgtcaagtg gaagtgagcc 86760
taactgcata gacttcactg ggctgttaag agaataaaat gaaataactg taaacagaag 86820
tgcctagtgc acatgcaaag gattattggg gctttctacc cttcagggat tagaagttga 86880
tagtaggcaa caagttataa gaaatacagt caattgtctg ctgaccaggg ctagagttaa 86940

ttgtctctgg aaaaaaggac ttgcctctct ttctcttctt cctccaaaac ttaagacgtt 87000
tgcagctgaa tccccaacag gattttgttt tcctttggga gagaggaaac agaccaatat 87060
acccccaaaa ctaaccccat aatttcattt cagcagtaaa gtgaggtcct tgataactgc 87120
cctgcccaac ctgcagggtg gttgggaaac tctgaatggt catgcatggg gaagcattgt 87180
gtccactgta aagagctctc cggagatgat aaatctcatc agaaggcttc atgcttgagg 87240
catggattct tggaaaaaca atcactctac gtatgtggtc agaatctaaa ggagatgctg 87300
gggagaggag ctaggtcagt ctccaaagtg gaacagtaga aactaatcat gtggagccta 87360
aacttatgaa ggtttttaaa atcagaattg gccaccttcc tttggaccat gagctcagat 87420
tgtgaggtgt gactaggtca cgtctccttc ctgcccctgt ttccctcctc tccctacctg 87480
tccctccttg accccaggaa aaattgccgg gatatgaaag ttaattatga cccaagggaa 87540
ttggtacaga tggggaagaa agaaatgcat tcaagagcat ttccatcagt attgaaatta 87600
cacagaaggc tggtgaattt gggctatcca ttcttgcctc cctctgtgcc cataattcct 87660
tggcctcctt caatttcatt ttccctttgg ttcagaggaa tgcttgatgg cttaagctag 87720
cctcagttgg ccaagcattg gagaaacaga gaggtgtatg acacagctac actcccatgg 87780
ggcttacagg gcaaggtgag agaagacaga agttgtatgt gctgggtgcc acgtggtagc 87840
tacaaactag aaatgagacc aggttcggaa gaggaagagg gcttgcagac ctgagtcatg 87900
gggacagttt cttcaggaaa tgggatctca gctctgcctt gtatgcaggg cttacataat 87960
aaatatgttt cattgttgtt gttgttattg ttgatttaat aagattttgt tttaagaaga 88020
ttttgtaaaa acaactgaac aaatgcaatc tcctgccaga gcaggcagca gcaaaggaga 88080
ttaggaatat aacccccttg gagacgttcc ttcacctacc tggtgctgga ttacctaaaa 88140
gcttcagcta agtagggtca ccccccccaag aaattatttt aaaaaaattg aaatctgata 88200
tttttagaaa atcttatcaa ggatatttaa ttggactatt tacacctatt tagggtcagt 88260
cggttttgga caagtatgca ggggtcttgg aatcagacca ctggggtcaa atcctagttc 88320
tgtcacttcc tagctgggtg accttggaca aagttacctg acttctaata gcttcagatt 88380
cctcatgggc aaaatagaaa tgctactagt acttaatagt gctctgagaa ggattcaatg 88440
agaaggatta aatgtatgta aagcacagtg tttgcccata ggaagctgtt atttataagg 88500
gaggggagca tcctaaggtc ctccgaattt aggagaacta aaaatcttac actgacttct 88560
cccttcaaca gcaccttcag aatctccttc atttttcata ctgttctttc aacccttga 88620
tgaatgagaa attaggcatt ctttccctgc agattttccc aaaccttctg ctttggccaa 88680
taaacatatt tttagtccca atcttgcatg ctcctttggg acttttcatc tgataaacat 88740
cccctcctgt gctcttgaat ccaataccct tcttccctgc cctccaccca gagtctcctt 88800
gtatctgctg ttaggcacaa tgatgacccc accaaggtca gacaatggct gtggcctcac 88860
ctggaccttg atgacccaca tagcctagag cccagagatc agccactgat ggaggcccag 88920
agggcagttg gaaaacttca caagacaatc cagcctgatt gttttgacat gcctgacttc 88980
aggctgctaa aaatgagctc gaggaatcag ataggaaaaa gagataggtg atgcaatttt 89040
attccatctc ccaatttttct gagtcaagag ttgtttgttt aactccagtt aaattagtat 89100
ttatccaaat ttcctgggtg cttgtccaaa gaaaagtacc ccagatctac aaattagaat 89160
ctgggactgg gacttaggaa ttggcacttt tacaattata ccagatgttt ctaatatgag 89220
tacttcaacc actaccctta tagaagtgct gcctaggacc ctctcttctg gcaggtgaag 89280
tggaaggagg ttttgtcgaa gggagattct ccacttcaac ttgagtgtct tggcttgtat 89340
ccgctttgtt tggttctatt tcaccaaagg ctttcatctt cacataaatt ttcttcagct 89400
ttaaataatt agttttggta accattggta tactggaaag aacattagat ttggagtcca 89460
ggtggcttga gttcaattct ctgctctgcc atttaccagc tgtgtgacat tgggcaagtt 89520
gccaacctat ctatgtcatt tcctcatgta aagataatcc cacttcacca ggccactttt 89580
gaggacccag tgaaatgatg tgtaaccatt ttaggaacac tggatcattc tacagtgcaa 89640
ttttttacat cagcttggag cctaccatgt aggcattcaa atccactgag tgtatggagc 89700
tccgtgcaca aataaaagga cttctctttt ctgcccgtgt acaactttgg tttccttaat 89760
caatagaatc catgacaatc ctgggccatg gtataaagat gggactttct tcctgtgaag 89820
gagtctggtc tgaacatctt ccaaactcca acataactga tgtcatttct ccacccaacc 89880
ccatttgctg tctcctgact caattgctag agaagccact taaggaaggt tcctggagtt 89940
aaggctgtgt ctgggccagt gtagcgagca gttttcaaca gtcagtcctc tttatcttct 90000
cttttcctgc gagcctttac taagcactgc ctcctcctgt ctccttactg catctcctga 90060
tggaatgcac aggtaaatct ccttggagag taccagccag gaacagtcca cagccaaggc 90120
caccgatcct caccgctgag ctccatcttt cctttcaagc tgtccttccc ctcccctccc 90180
caccatcacc atagcaacac agtggtataa aaaaatgaaa gcgctaaggc atctaaatat 90240
agtctgagta tcaactcttc cagcatggag ccgaaaacct agggaatgac agctagaggc 90300
atccagacga taactggcag ccaggagggt ggataagtca aaggaagggg tcaaggaaag 90360

aggggaagga aagggaacca tcacttgctg agcctgctgc ctgtgctttc tcatgtcacc 90420
cgcacgacaa cccaatgtga atgttatcat ctccaggtaa ctgctgaaga aacggaagct 90480
caaagaggta agagatttgg ccaaggtcac acagctataa gcagtagaac taagatttta 90540
actcaagttt ctatggcccc agaatttatg tgtttctctc tccataccac agggacaggt 90600
gcaagtgaga gattttgctg gaagcactgg gctttttgag caggccatat aaaaattctg 90660
agcccagagc tcaactaaat tattggaaga gactgggcca aatataaggc ttctatctaa 90720
gcagcacctg tgtttctcaa ggactgagga aaatgaaggg ggagggttgg caaggctgca 90780
tttcccaggg tgcgtgatta tatggcatgg gggtgggggc cattatgatg cccggacatg 90840
gaacttacac cagtgcagaa agggtgtgat tagaagccct aagccagaga atgttcagtg 90900
tgataaatgc cattattttt tccctcattc attcaataga ttttttttt agatggagtc 90960
tcactctgtc gcccaggctg gagtgcagtg gcaccatctc agctcacggt aacctctgcc 91020
tcctgggttc aagcaattct tgtggtccag cttcctgagt agctgggatt acagatgtgc 91080
accaccacgc ctggctgatt ttttttttt ttttttttt tgtatttttt agtagagaca 91140
gggtttcacc atgttggcca ggctggtctc gaactcctga ccccaagtga tccacccacc 91200
tccacatccc aaagtgctgg ggttacaggt gtgagctacc gtgcctagcc tcattcaaca 91260
gatattttta ttaagcatct gatgtgtgct taactctgga aatatagggg tgattagaac 91320
aaatgcagct cctgcccttg tagagcttat taggatagtg gagaagacaa ataaggaaac 91380
aattatacaa ttgattgatt ctttacaact gtaacatgta ctataagtac ataacagaag 91440
aatatcactt gcctgatgac ttcagtgaaa gggaaataca gaagttctta caaatcaaag 91500
caatcccctg ggccaattgt aaaggtgatg cccactttca aggtggacag agactgtgct 91560
agaagcttag cctcaaccat gggtttatat gattggtaga ccctgcagat ccattcccaa 91620
tggtgtatct tcatactaat catgaaatcc atctaatagc catacaagtg aggttttaaa 91680
acccaacaaa ctagactcaa atgaaatctg atgagggaat ttatgatttg ttcttcctac 91740
agcctttggt atcactgaca taaaactgaa tgtatgtgct gagggtgctt gtgtcttggt 91800
gatagacaag gtaggtggtc cagcccatgg tactggcagc ttaaagtcag ccagccatca 91860
gtgggaagtg cctgtgaatt atgcaggagt gggaggggag ggagtaggca gtaaagtaat 91920
gcatttctgt ggatccaaag ctttccaaac tacctgcaag tcagcaaata tgggggatgt 91980
tgtatgacta agtgagaatc agataatata atgtgtatgg agctctttag ttcttcagaa 92040
aaaaatgctg tctaaacaaa tagtgctgat atcaaagata atgatacagt accctaattt 92100
taatgctctg ctacctacct gccagctgtt tcccagggat gtggtaaaga tgaatgggca 92160
agatctggga aagtgttttg aaatccttga ttaaaggccc tccaggcaga tgtagaattt 92220
taaatgtgtt atattactgc cactattgtt atgctttctt ttatcacccc agaatttcac 92280
catctcctgt ttcaggtgaa cgagtctgcc tgactcttac ctgccctgaa tggcattgga 92340
aaggtagcag ccctgagatg tgccatataa acaaacatgt ttttaaccaa gggatcagga 92400
ggccttcctg gctggctcct gtcagctggt catcacctct ctataactct aggctttccc 92460
aagcttattt tatttccatc aataggacag gaatatgtaa atgtcctgct tgaaatgagt 92520
attggctaca agccatctgc ctctgaacag aggtgaaaag tggaaatcgg aggaagggca 92580
gatgtctttt gcaagggaaa cagactgttt tctgccactg cactctgccc aggcaaaaga 92640
gtaaaggaac agcactcagg agaattcact gaagcgaggg cagggtgcaa aaggaacttg 92700
agaaattggt actgggaccc aaaatcagat tctggcattt ctgggaaaag aaatgggcat 92760
gggtgggggt tttatctgtc aataaaagca tccagaatgg ggctagaagg aagtaaattc 92820
agttgccacc tctgcctact ggacagccac ggagaacttc tccttatcca aggtcgagga 92880
gccctccgga gtacatactg ataccattgg ttctcccaca catacccca tggagataaa 92940
aacaggaccc tggaagccct gtccgtgttt aaccaatggg attgaaacat ggaaatgaac 93000
tgccccacaa tccaccctgt gagagaccaa agagcagtgt tggattaaca gggaatgtta 93060
ccctgaaaag gcattcagct tccactgggg cagcaggtac agtgcaaaga tgatcccact 93120
taaattccta agacaggaaa taaggaaaga tgttgtggaa actcaagacc tctcaaagca 93180
tactcctttg tagttcttcc gcagaccaga ccacggaatt cagaaaacac cctacctggt 93240
tccaaaccag cacctgccaa acttctcacc ctcttctgac cctgtcctgg gagttaagaa 93300
aaaaaaaatc actttattgg ttgctccagt tataacttaa acagacagac catcatcaaa 93360
ttaagtgaca tgtacgactg cttattgtat gccagttact gtgctgtggg gttttggttc 93420
cattatctca tttaatcctc tcaaaaaccc tgttaggtag gttttattat tgcactcatc 93480
ttagattaag gaaactgagg ctcatagaga ttcggtaatt tgtcaaaagc cctaaaacat 93540
aattactgcc tccagatgtc tctgattcta aggcccaggc tcttaatcag taaatgatca 93600
aatgaataat gattttcatg gcatctgtca tcggaaagaa caatggagaa tatgcttaac 93660
caaagtcata accaaataaa tgaacttgac agcagagccg tgattctagc caagatgact 93720
attttcatgc atgttttgaa ggccaggaaa aggaggttag acttgtttgg gaagggaaac 93780

aggagctatc aaggtgaact tttcctaaga gtagcccaat aatagtgctc gggagggagt 93840
aatgtgtgca agaatagagt cagggagacc agccaagtgt gtgcctcagc atccctagca 93900
caaatcacac actaagcatt aagattgtct ctgcagtgag aaaggcctgg gaccaaattt 93960
gggctccacc acttactggt attcattaat cattcatgca ttcattcaac aaatatatat 94020
tgcgtgtggt ctatgtgcca gagactgtgc tgggtgctgg caaagaacac agacaaggtt 94080
cctgctctca tggagctttt attctgatga aggaaacaga ccacttacag ataaataaat 94140
aaacaagata aagggaaaca gatatgatgg agagtagctg gagggccaag cagaccgggc 94200
agacaaggtg gtggcatgta agctaagaca tttaaaaaga acctggtcat gagactatct 94260
ggagaaggaa agctccaggc agaggaagca ggtagtgcag aggccctgag gcaggaatga 94320
ggacaagata tttgagaaaa cagaacaaag gcaggcatga ccaggccgag tgggtggtgg 94380
aaaagtagta gaaggtgagt gggggagtgg gggcatcaag gtcaggcttt gcaggcttga 94440
tcagcgttct cactgtggtt ctggagccag cagcatcaat gttacctggg aacttgttag 94500
gaatgcaaat tctcaggccc cacccagacc tgctgagtca caaactctgg gatggggcac 94560
ctcattgtgt tttatcgagc cctccagatg attccgagta tgctaaagtt tcagaattcc 94620
taggttggat tatgcagttc aattttaatt ttaaatgcaa tgggaaccta tgaaagattt 94680
aagtagggga gcagcatgtt ataatttttct ttaaaaaatt gtttttaagc actcctgctg 94740
aggagagaat ggaccataac aggctaagag aaatggaagc agggagataa attaggtggt 94800
tattgcaaga ggccaggtaa gaagagaaag tggtttaagt agggtggtgt ggcagagaag 94860
acggttccaa gcagaggggg accacgctga caaataagcg cgggccactc acgcaagccc 94920
aacaaggcag aaggcagaag gcaaaagtga aggccagaga aaactggaca ccacctttcc 94980
agagcacagt tcaaaggcaa tgtcctcaaa gaagacactc caccctcctc ccatttcctc 95040
cctattgcct aaaaataaga aggatacgcg gcctatggca aaccttgggc aggcacgtgg 95100
gagctgagct cttgcaaagg gcagatagtt cctctggtga gagagaaaag gaagggccag 95160
tgaggagtga aggaagagac gaacagagag cccgaaaggc tgagaacgtt gtctggcttc 95220
ctgaaaggct taaggggtta gctctggagg gtgaactaaa agccctagtt atattaaaca 95280
cacacgcaca cacgcacgca cacacatgcg cgcacacaca cacacacata cacacagttg 95340
aaggagacct gcagtttcca aaaacaagag ttgtattttt tttgttcata tcatgaccca 95400
taacaatctc aaaagagaaa caatctcttg tcttccttgt ttaggcttag gagaacctgt 95460
agtaagtaag cagcagcagc ggaactcaaa ctcgactctt cctactgtca ttctctctat 95520
tacaccacaa ggcatcagag gaccactaga gtcgcctccc tagggttagg gttagggcaa 95580
ggtaaatgaa gtgagtcagc aagggcagga taggaacctg tctttattaa cattttgata 95640
ttttgtttat catggatttg ttgcattaat tgcaactttt aaaaatcatt gcattaaaat 95700
attattgatc ttgattactg agtttttagg tgtaccctta aatgttgcac ctctgactta 95760
ctagtctcac cctgatccct gtcctggatc tatgcctgtc tgttctatat cagcctcttg 95820
ctttgaccat aagaataact tcagaccttt aagcatagag gaaataggat ttctgtctcc 95880
cttccccacc tttgtgataa tctcagcttc tgcttttaaa gtctatctcc caagtagttt 95940
gcctactatg ttcctcccaa ggtcactagg ttctgtgaaa ctagcagcag gctagattgt 96000
cacattagca caaaggatcc actattcctg cagccgagct gggacaagca cttaggccca 96060
ctgactccaa cccttcaata gcctgggacc tacgttgtct ccaggtggta taaaacaaga 96120
atttcccctt tgactgggag aaaaagggaa gaactctaaa ttggaaaaca ggtcatctcg 96180
aattctcaca ggtggaaatt tctgacaacc cctttgggac ccacaattca acacacccca 96240
aatggggaca gtagctaaca tgcaacctgt aggctgttct gtcatccagt gccactgtgc 96300
tgcacaccac caggggcag cattctcatt ggcttctatg tgcctggagc ccagtgcagt 96360
tgtgcaacac tgcagctttg ctttagtgta gtccctgatg ggttcagtca agaaaatgtc 96420
tatagaatca gctaatctcc catgcagtta agtctctaat tgaaatattt tctctgctca 96480
gcccagggac agcaatcttt cctggatttg ctatttacaa ggatctctag aaattatcca 96540
ccagaaatat gggctttctc agagcttgag tggacaggga attaaggtgg aaggcagggc 96600
gttttgactg catttgaccc aagtcctgaa gagccagctc ctctctcttc ctaattatta 96660
gaaggttttg tttggaccca gtgtttcacg tgtatacaat acaaacttct ctcttttcta 96720
cttggatcaa atttgttctc tcaaaataag attcccagca gtgagagaag acaagacaga 96780
gagatccaac atctctaaag ccatgaatca gataaccagc cacttgttct cttcagtgct 96840
gggaacagat acactgttaa ataaaatgat tttatagatt cttctcactg cctttccaag 96900
aaggggattt atcaacttca gggcacagca atcatttatt cccagactac tggcatgcat 96960
atatatatat atttacttct cttgacttag aaaaaagaga gaattggagt tgtgaatatt 97020
cctgtctccc tcaccccagc ccccttgaag tgagtcagga caaacttggg gcccaaatgg 97080
agctgtaagt aactgagtca catgcagaga tgaaaccttc acagacccac tgatatggag 97140
gttgaagatt aaattcccct ttgagaataa ctgggtaaca ctcatacaga gactactttc 97200

aagaaggcca gatcctccct ctaatgtata gtgcaacgtt cctaaccctc agcccactcc 97260
gtcatacccc cactcacatg aatacacaca taagcagtaa tataaagcac ttcccaccat 97320
agggcagcaa agaaggaggg aaatctttat tatggaagag tggaaggaag gaagggaagg 97380
gaagggaagg gaagggtaag aggaagaatt ctcagggtga gcagaggaat gacatgtttg 97440
gggcataatg aagataattg aagtgcagag tttgtatgga aaaatttgaa aatatcaggt 97500
ggcaggccag gcatggtagc tcatgcctgt aatcccagca ctttgggagg ccaaagcagg 97560
cggatcacct gaggtcacga gtttgagact agccgggcca acatggcaaa accccatctc 97620
gactaaaaat acaaaaatta gctgggttta gtggcgcatg cctgtaatcc cagctactcg 97680
ggaggctgag gcaggagaat catttgagcc tgggaggcaa aggttgcagt gagtcgagat 97740
catgctacta cacttcagcc tgggtgagag agctttcttt tttttctctc acaaaaaaag 97800
aaaagttcag gttgcagaga tggatggatg gatggatgga tggatggatg gacggataga 97860
tagacattac agagagtttc caattcttag gatgaattgg aatccttaag tctttattct 97920
gtaagaaagg aaggggagaa taaaattttg tgattttaaa atattttcta ccctgtagag 97980
ctaccctaca aggcatgaaa accttaaaaa aaaaggcatc tactttaaaa gaataatgtc 98040
taaaaaatta gaaattccct ctttttgccc tgacctttgg gaaacagagt gagtgatcct 98100
tttgaggttt ttggcactgc cttgcctgtg atcatatcct gaaccctagg tccataatca 98160
tgcagttacc tcagatgtcc ctttccctct agccacaggt aacacgctct ccaggcactg 98220
ggaaagtggg taattaggaa agcagaggag tacccatggg ctgtgatgcc cagttataaa 98280
cccagacatt tcagaattaa cagaatgagc atcaagtcct caaatgggtc tacatccata 98340
aacatgtcca gcagtcagct ctttactgtc agtagagaca aaatgttcct acactttccc 98400
taggggaagc cacatcctca gtaggttatc tctgatgagt ccagctagtc acaggtatgt 98460
agaagctgca tgcagcagag ggctcaaagg agggtccaga atagatacca aagcaaaagg 98520
ggagtctgtg cacgttctca cacgcacccc gaaacactct ttttgttcac aaaatagatg 98580
gtgtagggta gttccaagag atcatttagc tcaggttcct gcctccataa aataaataag 98640
ccttccatat tagttgtctg ttgctgtgta gcaaattgtc agaaacgtag aggcttaaag 98700
caatacccat ttattatctc gcaagttctg tatctcagaa gtccaggcag gcttgactgg 98760
gttctctgtc caagttctcg tgagactgaa atcaaggtgt tggccaggct gggatcttat 98820
ctggaggctc tgaggacata tacgcttcca accttattca ggccatcagc agaatcccgt 98880
ctcttgtggc ttgaggttgg aggtccccgt ttccttgctg gctgtcatcc agggaccact 98940
ctttgcacct acaggctgcc tatgttccta ttcacaagac accgttcatc ttcaaaccaa 99000
agcagcatgt agaatctttc ttgtggctcg tatctttctg gctttccctt cttctttagc 99060
cagagaaagt tctttgcttt taagcgttca tgcgattcaa tcaggcccac ctggataatg 99120
tccctatttt aaaggtaact gtgataccgt ataacatttc aggagtgata acagcacatt 99180
tacaggttcc aaggattggg gcagaacatc tttgggggaa cattttagaa actctgcctc 99240
cccactcacc cataatcctt ttaaaaacca aatcttgaag cctttttttc ccaaaggcct 99300
ttttgaataa gcacatttat acctaacttc atcagacacc cactttgagc aaacactagc 99360
atgtggcaaa ataggctgta aatcaatcag aactattctt tcccaccaca atctttctca 99420
aacacattgg gagaatctga cactgtcagt ggtataccag agcagactcc taccatctca 99480
caagagctga ctgttaaatg tttagtaatt gtggacattg gttgttaaac tattagtagc 99540
ctgaaattga ctatagtgag agtattttca ccatggaaag caaccgttcc aaatcagggt 99600
ttctctttat tcctgggaag ctggtttatt agctcaccac tggctgtagt cctttagggg 99660
tcattacttg acctcctgta gcatgcagga atcctctcca tggccttttt tatgcatgga 99720
catcatccta ttttttaata ccaggaatgg ggtgatcact ctcttataag ctagttcatc 99780
tccctgatgg aatggtatgt ggtagagttg aaacccacct ccctggaact tcccaccaac 99840
ttcctttgga agcagcactt gtgacagccc cagaaccatt tggagtaagt agcatttcct 99900
ccaggagaca tctctcctct ggatccacaa atcaatagtt agatgcaaaa tctttagagc 99960
cacactgttt gaattcaatt cccagctctg ccacttattt agttataacc ttaggcaagt 100020
ctcttaactt ttctggtcct ctggttcttc atgtgtggga atggggataa aaatagcacc 100080
tacctcatag gttattatga atattaaatg agataatgtg tgcagagaaa atagcacctg 100140
gtctggcctc tacctatcta acaggttagt tgtgaggatt aaattactta atataagcaa 100200
aatgcttaga gctctgccta gcacaaaata agcactatgt aactattggt aagttaattt 100260
gaaatgtggt ttctagatct ctcatcatcc tagtcaccct actctggatg tactccaaag 100320
tccctctcaa gatatagtgt cagaattgac ctaattagtc cagcatttga ctgaaacgct 100380
agactttgac tccagccccc catccttgac tggcactagc attcaagccg cttctcctct 100440
ttccctgggt ctttaataga gtcagagcga cttctccagg ggatcttttg gccatggacc 100500
agtagcatcc acacacgctg gggccttgtt aaaaaggcag gctctcaggc cccaccccag 100560
atctactgaa tcagaatcca cacattaaca agatgcttgg gtgattcatg tgcacattaa 100620

agtttgagaa gcaccgcttt cagggacgag atgacacact tattttaaag agaacgccaa 100680
ttagagaccc taagccttct catggaacag gggccttccc ctcagacctt gggagagggg 100740
tcagggaaat atcagtgttg ggttgttggt gacaggtggc ggtgggggggt tcagtccacg 100800
ttcaaagagc cagaaacctg gcaggggaag agatggggca gtgacaccca accggaaaaa 100860
taaaggaaac tacaagaaga acccagctaa gagatgtgag gcttctgaaa gctcccatgg 100920
aaaggttcgc agctcctcca cctgctcggt ccagctgccc caggtcaagg aagctctgtg 100980
agtgttagct gacccggagc agcaaggata cattcagaag tgatgaaagg gaacgcttct 101040
tgacagggta aagagtcatt cagtaggaat gagacaggaa gaggtcacag agtcagaagc 101100
ccagcctgta ctcagagatt atttctggca tgggagggcc gaagggttag gaggccacct 101160
actcacaata caatacagag gcagatccac ttattacctg cctgtgctgc tgggatttca 101220
gtgtggaaat tctgtgcctc ctcactgtgg ctgcagcttg ggaatgacat ccagagctta 101280
cccacctgca taagaaataa gctataggtg taataggggg acataggcta aaatcctagc 101340
tcagctgctt aatagctgtg cgactgagca agttacttaa cctctttgag catctgtttt 101400
ctcatcttta aaatggaagt aatcataatt gaccaggccc agtggctcac acctataatc 101460
ccagcacctt ggaaggccga ggccagtgga ttgcttgagc ccaagagttt gagaccagca 101520
tggtgacacc tcgtctctag aaaaaataca aaaattagcc aggcatggtg gcaggtgcct 101580
gtagtcttag ctactcggta ggctgaggtg ggaagattat atgagcccgg gaggttgagg 101640
ctgtggtgag ccagattgtg ccactgcaat ctagcctgga gacagagtga gactgtgtct 101700
caaaaataaa taaataaaat aataatatct atgttaataa agcagaaata agaatgaaat 101760
aagaggcctg acatggtgac ttatgcctgt aatcccagca ctttgggagg tcaaggtgag 101820
aggatcactt gagcccagga gttcaagatc agcctgggca acttagtgag gtcccatctc 101880
taccaataat aatttttttaa aaattagctg ggcatggtgg catgcacccg tggccccagc 101940
tactcaagag gctgaggcag gaggacggcc tgagcacagg agttgaggct gcagtgagtc 102000
atgatcacac cactgcactc cggcctgggt gacagagtga gaccctgtct caataaataa 102060
ataagaagaa tgaaacaaga aagttcttct tatggttctc atggtggtga gcacaatgta 102120
agcatatata ttatcttaga attcttcctt cctgtataaa gaaggcctcc tccaatgtat 102180
taatcatctg ttcaactaat aaatgctgct tactcccact ttcactctaa aggaactcaa 102240
tggctaaaga gaacccttcc cctttgcagc accctgagga tcagaggcct gatttgaatg 102300
tcctcgatgc aaaggactat ttcaaaaggc cagccaggca gcccagacat gtatttccta 102360
atcgtctcca ggttgtttga tagaagatct cctgggagca ggtttccgca gcagctcagc 102420
caggtctgtt ctgggaacgc tgtgtgcatt ggcacctccc ttggcagaaa gcttggagga 102480
aaggcaggtg caggtcctgg agcctctgac agcattactg gctctaggag tagctgctca 102540
ggataatctg tccccatgac cattaagtaa ctgccactgt gcgggaagaa gaactggaaa 102600
tgggggggccc aaaaaaatct gaaaaccctc acttgaacca gtaagttata ccctgggttg 102660
ctgttggaga gagcttcctt ggagtagaca aatgtggtat gttaagtaaa ctggggatct 102720
aggtttgatg atactgggtc tgcagcttct ttgtcccact gaaaatcctc gggcattcca 102780
tgaaagtagc cttcaaaata tttttgtctc taatgacata tttttgctgc aaaaagatga 102840
gtggattcat tttacgaagt ctcaagtgtg ttagaaattc accatgagtc actcagcaag 102900
ttatgtttga gggcgttctg tatgccaggc actgtgctgg gcactgggac tactgtagca 102960
agtcagatag acaagaactt gcttgatctt ggaagtaagc agggtggggt ctggttagtc 103020
cttgaattgg agactgcctg gagatactgg atgctgcaag cttttgaaaa aagacaagtt 103080
ctctgtactt gcagagctta catccagtaa ctaactaact aacttcaggc tgtgttgagt 103140
gactgaaagt ggtggagcca ggagtcctct agataaggta gccatggaag gcctctccga 103200
agaggtgata agtttactca gagacgcaaa cgatcaggat aagcacagac cccggtgaag 103260
agcgtcccag gcagagggga tagcaagggg attgccctta ggtgggaaag ggcttgattt 103320
gaggactggg aagaccagtg tgtctaggac acataagcaa ggggaggacg ttatgaacga 103380
ggtctgaggg gtcagcagcg actggatcat gcaagctccc ataggccatg gtaagggctc 103440
tgtgtgtact acaattacag gatgcatgat aggacctggg ctgcattttt aatagttaac 103500
cctggctata atgtggggaa gggattgaag aaagagggca aaggcaggaa caggaaaatc 103560
tcttaggagg ctactgcaaa gcccaaggga gaggtgatgg tgttttgttg ttgttgttgt 103620
ttgttttgtt ttgctttgag aaggagtctc actctgtcgc ccaggctgga gtgcaatggc 103680
acaatctcgg ctcactgcaa cctccgcctc ttgggttcaa gcaattctcc tgcctcagtc 103740
tcccaagtag ctgggattac aggcatgcac caccatggct ggctaatttt tgtatttta 103800
gtagagacag agtttcccca tgttggtcag gctggtcttg agctcctgac ctcaagcgat 103860
ccacccgcct cggccttcca aagcactggg attacaggtg tgaggcaccg cgctggccaa 103920
atgatggtgt tttgatctgg gtcttaaagg cagaaggaag gggggtagta aattaactgt 103980
gctggggaag agagggaggc ctgagagtga ggaaagaatg aggggtgatt ccaggtttag 104040

gaaaactggg caatttgtta gatgatggtg ccattgacag aaatgggaaa gaacaagttt 104100
ggaaagaaaa ctcaagatct ggctggtgac ttgtattaaa cttaaagcct catttgtgac 104160
ttgagcagaa gtaaggactt tctccagtgt tcaagagctg gaagggattt ttctagcctc 104220
caggcaaggt aataccataa gtcccaacag tgatgccctc cctgggaatg atctcaatgg 104280
gagaatccta taccctgcct cctccattca ttccttgctc tgatggtggt tctggctggc 104340
taacctaagt tactcttgcc actagttaac gcctgtcctt atttctcttg tccccaccta 104400
agatgtcaat caaaacagca cgagccatgc tatgtcacat gacatgttgt ctgtccagcc 104460
cagagcttgt tgctgatggg ggcacagact agattttgag agaaatctct ctgttaccac 104520
ccttaacatt ccaacccccc ctaatagccc atttaggatt tatcatactg tttcatccaa 104580
acctttcatg acctgatttc tatttccagc ttcaaccacc ccttgggtca ccacctgtac 104640
ttattgagtt tcccctagttt tctgaattaa tgactgaaga tgataagctt cccttacata 104700
tgactctcaa accaccaaac tgggattgtt gttactctta gtgataatgg ttgctattta 104760
tgaaactttt aatagggaac acaaaccctg cccagaaatt catataaatt atttcattta 104820
agaacatcac aaagtaggtg ctattatttg accttacacg tgagacttga agaactttag 104880
agcattgccc aaggtcaccc agctagtgag gggtggaggc gggatttgaa tccagctcat 104940
ctgtctccat tacctggaag aaggaaggcc agagcatcat ggcctttcac aagttgaaga 105000
gccacgggct ttctacggta gccagccacg cttttccatg actggggtgg gtgtggcaag 105060
tgatgagggt ttggagttca tgtggtgggg tggcagggac caggtgtctt ggtaactgct 105120
gttgcattca cttcaggagc aaaggaccag atctgattct gcaggatcaa caatatggac 105180
actgcaggct ctgtagacat ccaaagctct aatggtgact tggggaagct caggagggca 105240
gggaggttgt acccatttag aatgtaaaga ttcctatttt ataaaaaaga aaaaaaggag 105300
actgaaggcc tcagtctcct ccaacaaagc caggctgtgg ggtagcagag tctcaaaggg 105360
tgcaggccca tggccactgc ccagggctcc tgctcaggcc tcctcactcc cacaactgag 105420
gggagaccca gttccacacc cacccaccta gcagtgtctc acacccaccg ggagaggtct 105480
aaacatcttc cctgggaaat ggtcccaaaa tgtccctgca gtaagcaacc atctggagag 105540
gcccaggtct acatctgttt ttaaagctcc aataaataaa taaatgaagg aagaaaaaaa 105600
gaagaagaaa tgcagaacag ggtgactaaa attggcatgt atttttaaat gtttatatta 105660
acaaactaac acctttttaac atgaaaagca atataattgt gctagccaca aaatcatcgt 105720
aggactgaga aaggaatcgt gattctgaga gccctagagt taatgtgatc cagctggctc 105780
atccctgtga ctgcagaagc ctgtttggag atagtgtcag tagcttttca ggccctctgt 105840
gaattgccag aatgtgtgac atgagccaaa tttcccccca gcatccccgc cgccgccacc 105900
accacccccg acccaaccct cccgccggct cccatagaat agtcactgcc atacagaaaa 105960
agagaagttc tactatttct gggcaagatt tccacaaacc agtttgtccc tttctgcttt 106020
catgaaataa accatttgga tcaacgtcag ctgattgcaa aaattttccc ttgtctcaaa 106080
agcaagactg ataaggaagc aaacatggga ggaccttagt ggccgagcct ttatgtgtat 106140
gttatttcat tgctctcata actgccctgg gatgctgtaa gcatgattca tcctgtttgt 106200
ttatcagtta aattatgtat ccaagattac acagcctatc caggattaga actcagagcc 106260
ctcggctgtg aagcttgagc tctttctttt cagtcttcaa atatgatcat gccatgaagc 106320
agcacaaagc ccaggaggag cccagtgagg ctggaggggt ccactggcag ccactctcct 106380
ccgtgccccc gtggtgttgg ggcaaacttg gatctttctg aatcttttaa ctgtttcctt 106440
ctcttcccgt ttttgtctgc tggctgactt gtcctacact ctactccttg cttatgatac 106500
ttatttttcc atccacagca aaacaattca catcaaggta attgatgatg aggcatatga 106560
gaaaaacaag aattacttca ttgagatgat gggccccccgc atggtggata tgagttttca 106620
gaaaggtgta gtaccctgtc ctccacacta acactaacat tcttctctcc tcttctgttt 106680
cttcctctcc aacccatttg tctcctcctc ctcttgtctt ccacctctct ggttcccttt 106740
cccttgtctc ctctcttgct ctctctcctg ctctcttttc actcctccct ctcctctgtc 106800
ctctctctgc ccccagctct gtcctaacac ctgccagcct gacacatggc atccatacga 106860
gggatgctca agaccgatgg taattgttct gggataagga aatgagtatg gggaaagaaa 106920
gagccaaaat gctggagtat catgtgcggc tcttggcttc tccagaatgg ctgggcataa 106980
agggggggaaa agggaccaca tagcccagca ccagacagaa gagcagcact gagaaacagg 107040
ctttcagcac aaatttccat ggggcagtta ttctcagggc taaacttaga gtcccaggaa 107100
gttgagaatc aatgtatttg gattacagtt cattcccctc ccaaaagcag gctttaggag 107160
ccaccttatc tgccatgttg ctactatcaa gacttgtttc tcctcctgac cttgaggaag 107220
ctgaaagtac aggtttgagt tccagatcta ggtcaaatat ccatttgtct tcctatgttt 107280
ttcctattaa gaacacccag gtgtggaggc agagagttag aatagtggtg gagatcatcc 107340
tgacccaaat ggaagcttcc ccaagaggtc catggggctt ctcagagtgg atggaatctt 107400
tgccttcaac ttcaatgacc ccatacatcc catggcctcc aatagacaag tcaagaagtc 107460

ctttcctgaa tagatcatac tgtggagcag ggagctgcca gtactgaggg caatgttcct 107520
tccccttcca agctgtccct catgccctcc agtacatgcc tgttgtcaca gagcacccca 107580
atcccatccc acagcagagt tcctgcagca gagaaacagg ctcacacctt gtagacagcc 107640
ctggggtccc atatctaggg ccaacagaaa tattcccaaa aaaatgcctc ttgacaatca 107700
atgagctttc tcttttgtcc gctgagcaag gtataaaaag atgtcaaaag aagtacccaa 107760
aaaggtaata aaaatgtaca gtcgtgcatc acttagcaat aaggatacat tctgaggaag 107820
gtgtccttaa gcaatttgt catcgtggga aaattataga gtgtactttc acaaacctag 107880
atggtgtagc ctacaacaca cctggactat gtgggcctat tgctcctagg ctacaaacct 107940
gtacagcatg tgcttgtact gaatattgca ggcaactgta gcacaatggt atttgtgtat 108000
ctaaacacat ctagacatag aaaaggcaca gtaaaaatat cgtagtatat agccttatgg 108060
gaccactatt gtagatgtgg tctgtcattg agcaaaacgt ttttatgtag catgtgactg 108120
tacttgtaaa gtacacacac cacaaatgca cagcaagtcc tgtgccctac aagccccttt 108180
gggtcagtct actacattat aaatggcaaa gccgagcacg cccacagaag gtagcaggaa 108240
catcagagga tctgaagaga catttaggta aatgctcttt acccctttaga gcatttagtt 108300
cttaggcctc ccctccccca atctcccccc cgccccccgc caaaaagaaa aagaaaaaga 108360
aagcagaaaa ttacaattct ggctcactag taggacctgc tagccaccat tgtgattcca 108420
tgaaggacca gaagaaacca tataggaaga atcaggccca cacggcaacc tctccacatg 108480
acaaagagcc agtctttgga gggcagtgaa tttcaaggaa agttttcttc cctgggtgac 108540
ttgtttttaa aagatgttat gttttgttga gatacccaga gatgaacaga aacttccatc 108600
accttgtgcc ccagacccat gataattcac attgaggaaa ccagttttgg aacacatcac 108660
ccctaagtga tagaagccca aaggtgattt agaatttgat gatttacatc attttcttca 108720
cattttccca gaaatgcatc agctgtaaat agtaaaggat tcctatgtaa tattgtggtt 108780
aatacatatt tattttagtt cccaccactg aagccctatg agataaagaa tgagaaagat 108840
cacacaattc tacctccctt tcttctctct ctctctctct ttctctttct ctctcactct 108900
ctctctctct ctctcttctc ttcctctgtc tggttttcct tcctcataaa tacttttctt 108960
ttaaaatttt ctttctgaaa ctcacaatgg aagtgagtat agacataaag aagggacaca 109020
agccctgggt tctgttgaca tattccctgt tgtgggaaga ccctgggtta ttcccagtgg 109080
gttagtagtt tacctgttgc ccagagaaat gccactgtta tcatgtgaca cccagtggaa 109140
tgtgctgcct gactcacttc ctactaactg ttggcaaggt ctaaaatgac tcctcctcac 109200
cattacccgc cttctgcctt ctcctccccc tctgtccttc tggctccctt cctttgccca 109260
cctttccttg cctcctggct ccctgccccc tcacccgtaa gaacaactat gaccaagaag 109320
acaagaaaaa ctaagaccat ttattacctg agaacaacac aatccaccat ggtcctgttg 109380
aaagccacca tggtgggact ggactgcatg tgccaggaat gacggggaat gattttaaag 109440
gctgtgctcc aggtgaccaa ccaatctacc gacccagtcg acacactctc tctcttgttg 109500
tccctacagg aaaaccataa gggttaaaat agtagatgag gaggaatacg aaaggcaaga 109560
gaatttcttc attgcccttg gtgaaccgaa atggatggaa cgtggaatat caggtgtgag 109620
attctttaaa aacaaaacaa caaaaaaaaa gaaagaaaaa ttaaaacaaa ctgaaaaaca 109680
acaacaaaaa agaaaaagca gctatatttt tgtctccctc cttttcttcc cttctcctcc 109740
tttctctttt tgaccaatgg attttttttat tcttttccct cctgtattct cgctctcacc 109800
ctgtttcggt atcatctctg ccttcttagc cttagcttat tccaaattcc tcctttaccg 109860
ccttctgggc agcactgcag cctcaactcc tcattaccct aatgagttat ttccctgttt 109920
tgctacaatt ttcaattatt caattgccat gggcccctgc actctccccc accccacccc 109980
tacactgtaa cctgtaaatg tgaaaattcc ttggtgggtg gggaggagaa gaaaaaaaag 110040
gaatgtgatg cgatgcatgc ctgtgccccc tcctgccttc ctcccctgcc acccctcact 110100
ctttagcctg gattgaatgt gggggggtct gggatggggg ttggggcctg ggttgcaatg 110160
atgctttgac agttttctgc tgcattcccc aacttccttt gaacgcttgg caggttattc 110220
acttgtggag tggcccatag gcccctctgc ccttcgagga ggtaagtgta ttttctggct 110280
gtttcacagt tgggcagacc gtggcatggg aaagtgtacc aattgtcaga agccacggct 110340
tctgagagct ctgagagaga gagttgactt ctggggtaat catgcaatct ggaattctga 110400
gctattcttc ctcctctggg catcccaccc catgccattc tatgttccta gcccaaggtt 110460
gggtgcctca ttcaggctac tttgggacaa tgcaacctct aaagcagaaa attgagagtt 110520
cctgaaggga aggaaatagt tccaggtatg aaaattcccg tagccagggg ccccagaaaa 110580
ggactgacat tgggcaggcc tggagtgttg acttgtggat tttccaacag aagagactct 110640
aaatgatgca gttggtgctg atccctgaca gacaggtgtt ggaaaggtca cagatgtctg 110700
cctttgcttg gcatctgcaa gagaaagtac cgcccagatc ccaagatagc cctcatccca 110760
cactagagaa gtggcctcat ctcctgcttt cctcaggacc tgcatctgag aatacctgcc 110820
aggggctcat ccctaaagga ctgattatgt tgcaaccagg gtagaagtaa ggaaggattt 110880

cttcccttga agaaaatgat tggaagccac tactttgaat ggcttccaat catttggagg 110940
catagatgtg ggaatgggtt agggtgctcc tgggaaataa caagaggacg ttcacactcc 111000
cattcaggag agatatgctg ctgggagcct cctagcaaat gaagcagtga aatccacctg 111060
tttgtcaaaa aggggtgatc atactgcaat tagttcatat tcatgtgaca aagagcagca 111120
taaaactttc cacacgagga cagagctaag agattcagca acaacattcc caaaggattc 111180
tctacaggcc ttctcagtgt gattggtcat ttctcattgt ctgctgggga ctctcctgca 111240
gagctgacca cttctgtgcc tgcgctggtt tggacacacc tgatgctcta ggggcagaac 111300
tcctctcctt cttcactgct ggttctcttc gtcaccactc aataaaacgt tgccctcagc 111360
ctgactgcca aaaagtgctg gaagaaagaa attatctctg gttctattgt ttcccacatt 111420
gtattcttgc ccaacttcca gttcttgcca ccaacaatat tctcagaggt tgcctcagca 111480
cctgccctac ctcattccca cctcccttga gcatttattc catgtattca taattggttg 111540
gaagcagcag atacccaagg ccaattgtaa gtcaccttca tcagtttcca cagtccaagc 111600
tacttagatg caaacgaaag cagcacatgt acagcgtaca ggaaggaagg cagtggttcc 111660
agacaagagg aagagattgg aagtccatac atgcctttat tccaccagta aaaaggctct 111720
tctcttatgc ctcccttaaa acctctacca acagcaggac agagagtgac ccaagataag 111780
tcttcaagag acctaaccaa atgcaaatgt ctttggctaa tccccattta aggacatctt 111840
cctgttttgc acagattctt tgcccaagga aatgtcagca atgccctcgt ggagggagta 111900
ggtgagaaga caaggatttc agcaagctat ctgtgtggtg tgcccccaga tctccccagt 111960
gaccgagatg ccaagatgaa gagtgccaag aagaaattgg tcaattttcc agctgcctat 112020
tttattgtct atgttttcta ggcggttaat ttccagtttc ttcagtactt cccgtatttt 112080
gacattagac cataaggtga aaggtcataa aacctgattg tctagactca gaagcaaatg 112140
gaaacccatc caaatttcca gaattccctg ctgttctcag agtgagaaac agaacagtgg 112200
aaattgcttt tcattatcac tactgcatgg gagagtctga aacattcaga atggcatagt 112260
ctttgcatgg tcaaaatgac aattgcatta aaaaaatgag agactggatt tgaaatagga 112320
gactctattt ttggcaaaca aaacagactt cagagttgag attaaaagct ctggatgagc 112380
tgggggatgg aaaaaaggga aggaaaaaag ggagactgaa taggaaacac agttgctctg 112440
gagtctagaa gtggacttcc gagagcaaca ctgagcaaca taatcaagac tgttgggcct 112500
gggcctggac attggaagcc ttcggataga aaggaaagct ctctgtctct ctctctctct 112560
ctgaagaatg gggcctgttt ggtcctcctt tttcgacaac cgtgggctca tcttgacaag 112620
ctgcccagat gcttcctaat tactcacagt cctatgctct ttccagcttg tccctggggt 112680
gtctgagcag gaataaatga ctctcacctg acccagggga tcaatacagg ggaaagttca 112740
gctccagctt ctctcatgag cagcagcagg aaaaacaccc tcgaggtatt gtgtcagtca 112800
aagctggcct acccaggtct tgctgaccca tctataactg ctgagcagaa agtcttggat 112860
tcatggagac aatgaccaga gaatgatgga attccagcca actgcaggcc ttctcactac 112920
tctagggatg ggccagatgt tcggtggcat gtatgagtga aaaccagggc atcagggacc 112980
tttctggaag agctgccttt gtctgaccca cctgtgttca tttatgtgct gggatctctg 113040
atctcccctg gaacttgggg gaagctcttc cacgcaaact cccggaagga gcagaataaa 113100
caagctcttg cctatctatc tatctatcta tctatctatc tatctatcta tctatctacc 113160
tatctgccta tctatatcta tctatctcaa tgtagtgagg aaagccattg atccattaac 113220
ctttggaatt ctacatggga gatacctaaa aaagtgaact gccttgttta tgtatcatgc 113280
agactctgga tccacatata tctcagtggc tgtgaatata ggatgattga tcacaggcct 113340
gagttgcatt cctacagatt cttaggaaaa aaattgattc acagacatgt ccccccctggt 113400
tcccccacaa cacacactcc ttcctcagca atctctatca gtcaccaact acacgttgaa 113460
tatgtggcaa gctcttccca gacctttatc tgagagccaa ggagtgaggg gctgtactaa 113520
gatatcatag aaatgaaaat gtggtgtgtc acaagtttcc ttaattctta gatcttaaac 113580
tctaagaggg ttcagcataa gtacaaattc aagggctaga gacaacctgt attgggtgtg 113640
tctttaactc agtttcccaa tccacatagg gaccttgcat ttgtcatctc tcatctatgt 113700
atagctgttg gtatgacagt ttctctgttc cagaatacct gaactctgac ttagcctgtc 113760
ctttctgaaa cagaaaaatc acccaaccag agatctatga gatctatgga aaagacagtt 113820
gccaaaatag acagcaaaca gccaaactta attgaacact accacatgca gggactttgc 113880
taagcagagg tgatacaaaa tgggaggagc ccatagccct aacttccagg atatatctac 113940
ggtaaagaca aaccattcaa ggaaaacatt ctgcaggact tacctttttg ctaagtcatt 114000
cttttagggg aaatcaaagt tctagtcaac gtggcagcta ggaaggcatt tgtggtgatg 114060
gaaaccttat gagcactgag aagctgagca tgagttcagc taagtcgtta gggatggaag 114120
acatagacct gggcactgtt ccactcttgc acaatgctac ccatttcctt gagctcccat 114180
tcaagcccca tggtcatttt tgccactcat aagttagcta ctctggcagg gttgcaactt 114240
acacagtttt catgataact ggattctcac tccttttttt acagaatgga tgtgataacc 114300

tggtatccta cacagtcatg agtgaccaac ctacccattt ggttccccat cctcattcct 114360
ccattcctag ccctagggta gccgggaaag cataggagca aatgccctta ccagggccct 114420
ggtgctcagc agcctctccg gctgctcaca cctcttgctg ctgctctgtg catgctccaa 114480
aggctgcttt ttgcgtatgg ctgctgagct ctcacctact aagctctctg ctttccttat 114540
gctgccagca accacaaaac ctggtgatac tttcaagatg ggacattaat gctctttcct 114600
tttctttctt ccatttttct ggtatccatt tgcaaacagc gctcctgtta tctccaggta 114660
agaggtgtct tgtccccctc ttttctttcc acttcttgcc agtgccatta tttggtttaa 114720
gaccaatgtc ctttgattta ttgaataaga actgcaggct caagttaacc tgacaatttc 114780
tcccaaggac tgggagattt attttcccac atgaagcaat tatgagaaag caattgtgag 114840
gaaggcaatt ccttgagcat cacttctgtc tggggacgtg ggttaaggca tagctgatcc 114900
tctctgggac caggaagaga aattaagctt aacaaggaga tggtgggtca tagacttctc 114960
ctgagtctta attcatctgc catctcatgt tgtgggggaa gagacagtga gattcagagc 115020
tggaatctcc taatataatt gtgacaggat ttgaaaaaaa aatactttaa tcccaaggga 115080
tccaggaaat aaccaaacct gttgtgagaa taggaaatgc aatttttaaa gaatctggaa 115140
ttttaccagt cctggagatc ttccatctca tcacagctga gacttaaatt gctagaattt 115200
tggttcattt gtcattgacc cttaaagtcc tatgtgccgt gaacaagatg aattaggatg 115260
ggggattggg gcagtgttct ggctggaaat ataaatttta gagaatttat tttgaagaga 115320
ttctcatgca gaatctaggt gctatagagg acgtacacct actttgagag tatgcttgca 115380
tgagtggaaa ccaatcataa acaacattca acttcatgag cagatatgaa agcattttca 115440
gcatatctag caatactata actctttgtg caagcagagt ggcctacaca agacagtttc 115500
aatatatttt aaaagaacgt cttacatttc atcagtcctt tgaacacaga aaaaaatgtt 115560
aaggccactt aagaggcaaa acatcttaca gagttcattg atattcaaag tcacctacag 115620
gctacatctt gggttcagga aggggcggtg tacatagtaa ggacatacgc cttctgggag 115680
ccttaaacaa acaaaaaaaa tgtaggtaac tcctacattt ttcttttgtg gaaaaaacac 115740
agttactcca gcttccttgg ctttttgctt ctttttttata ccaacaaaat aagggctatc 115800
ctcaaccctc tgttcttcat tcttctccca gggtattgat ttcataacat tgggtttttc 115860
ttctctactt cactcatcct cttgcctgtg aaggtatgta aggcttcttt gttccaactc 115920
tttcctccac ccgccccccc tcacataaat gcataacaaa gattgtgatt taatttaagt 115980
ttctttctac ttttaacata tttgcaaaca tcaatagaag ctaaaatggg aaaaaggaaa 116040
tgtttctttt cctagctctt tcaatctgta agcctttaat ttaggagcgc tgattagcct 116100
ttcaattcgt tggaaatctc aaatactggt tttaattttc ctaggtggac agagacagag 116160
ggaatatgtt cattctgagc taaccacccc cccaccccca agctcaggcg ccttgcagga 116220
agagcactag ctacatcact ctgcagagtg ttcacaacat cctattcttg tctggcctgg 116280
caagctcttt gtccttccaa tatttgttca atcttccatc ctattcatat tctatctttc 116340
tctcccctcc cagcctctct tcctgttcct agaactgaga gtttatttag tcagtctgaa 116400
tatctagatc acctgccatt tattctcttt acttgaaatt ctgaggagtc acataaacaa 116460
gatatcagaa tcactatggt cctctaaatt gaagacttat aattctctca agaaattaac 116520
aacatttgaa tttaaaggaa agatcatgac aaaaatagaa aaaggcagga attattgcca 116580
aaccgagaaa ctagaaacta gaattaactt aaaggcatgt gactcaatca attaacaaat 116640
atatacagag agcctctgtg ggactgtggg agatccaaag atagaggatt ggttatttgt 116700
caaagggatt tttgcagaaa gctagatgga aaaactgact gtcaccacag aggtggacag 116760
gtcagtaagt agatcaatat cctgccagat ggatatagtg ctagattgat aggtagacaa 116820
ggggttagac aggtacattt atatgtcact ggagagctca ttatattggt ataaagttat 116880
tgtgtcacat gtaaagtatg acatggggga attggggagg aaggagtgga ataatactgt 116940
cgctgctaag ataggcattg tgatatggtg cttaaacctg caagtaaagg aaaagagtat 117000
ggaatctgtg tgtctttttc taagggcttt ttcccagagt agcttgcagt ctggcttcta 117060
gggttgctgg cctatagcca gaaccctaga ttcacccaga tttaccttca gaattaacta 117120
atcagagact caaattcaat agactaaatg aagtcaggct gctagaggat gtctgctgac 117180
ttggacatat gcagaaagac atggatcctt gagaaaacat tgtttccaaa agtggccacc 117240
agcactagag gaaggacagc accacggaca gctcccagac attttaggat tgccttctgt 117300
gtttggtgcc cgaacactga gcaaaacagc gaactcagga agtctccaca cactctcata 117360
ccatcttcat gcagtccaac taagaaaatt cttacataaa atataaggct gtctgcttgg 117420
taatttaaac ccttggctta tagtcttttc agtgaatttc tttccttgca aactcgagag 117480
ttggagtctc acgactgccc ttgcttcacc aattccccag ctagagacaa aagaccttct 117540
tggcctctga cccattttgt ccttgagatt atccaaggac tacaggattc ccctaggagg 117600
tttactgtgt ggaatgaaag caattaagga gctgaataaa agaaataatt gcatgtgaga 117660
atgtggactt ggatgggaag atgtttaaat gagctctgaa agaaacaagc tgccaagagc 117720

aattttctaa ttaaagggga ataaaaagat tcaatctcta tttcactcta atccagaaaa 117780
catgtcttca tggagaagtg ctcttaaaat ggactcatca gccaaagtgg aaaaacaaaa 117840
aacaaaaaaa ctgttcaaca tgagaaggga ccattggtaa atgagtcaag atgctgtgaa 117900
accagtagac atttcctttg aataaatgta cttctgcacc ttcaagaact cttacaggaa 117960
gtggttgaac aaacaggccc aaaagttcaa aatagttcaa ggtcaaaaca cttgcccttt 118020
cttcccagtt ccccaacatc tcactgagtg tcttgagaac ttcacttgat gctatttctc 118080
aggagatgtt taggtcaggt tgtccaccca ggtataaaag agaaagagga acgcttatcc 118140
cagtctgcaa ggcacattct catggtctgg ttataaagtg tttagtactt cataaaaaag 118200
gcactaaaaa tatatataaa ctccccattc ccaagagtta tttgctttgt acccactgcc 118260
catgcctaat actctgagct gtatccttcc agggaatgga aaaggtgtta aagcgagtct 118320
gattttgttt tgttgcagat gtgacagaca ggaagctgac tatggaagaa gaggaggcca 118380
agaggatagc agagatggga aagccagtat tgggtgaaca ccccaaacta gaagtcatca 118440
ttgaagagtc ctatgagttc aaggtcaggc aaacagtgag gtctaattga ataataaata 118500
aattaaagtg ggaggcagaa gacctggggt gttttttcc actttcacta gtgaatatgt 118560
gaagttgaaa ctgaacaaat cacttaccca ccccaggtct cagtttcccc atttgtaaca 118620
tgaaacaaat agtgctgacc atttgtatgc taggaatatt gttaggaaac ataatataga 118680
atgtgaaata agtggactag aaagtcctga gatgtattat cattattgtt taactgtgtt 118740
tttaaagcaa aaatattaaa actcactact acagggcaag atatattaac atcattatta 118800
ttattcatta ttgtattatt ctaaatagcc aatttcaaaa gtcacaacca ggccaggcag 118860
tgagggactc acgcctgtaa tctcagcact ttgagaggcc gagatggaag ggtcacttat 118920
acctaggaat ttgagaccag cctgggcaac atagggagac tccatctcta taaaaaataa 118980
aacaaaataa aaatcagctc agtgtggttg tacatgcctg tggtcccagc tactcaggag 119040
gctgaggtgg gaggatggct tgagcccagg aggttgaggt tgcaatgagc catgattgca 119100
ccactgcact ccagcctggg tgacaaagtg agaccctgtc tcaaacaaaa caaaacaaaa 119160
agattacaac caaaaacaaa gggaaataga aggattgcct caaaagagat cgcccaaggc 119220
cattccatgc gtaactgtca gaacaccttg gagacagggc atctttcatt cctttgaaga 119280
accagactcc tcattggttc tgagcattct aacctcatgg ttccaagttt ttctcttctt 119340
aacagactac ggtggacaaa ctgatcaaga agacaaacct ggccttggtt gtggggaccc 119400
attcctggag ggaccagttc atggaggcca tcaccgtcag tgcaggtgag aagtgtctca 119460
ggctggcctt gctgggagaa gcaggcaacc tctgagaagg aagcgtaaag ccacgttaac 119520
agcctgccag tccctaggaa ggcttgtgtg ttcagtcttc ccagctctgg tcctaggtgc 119580
ctgcttggaa aagaatcatg gcgtatctga aaaacatggt tatctctggt ttcaaatcgt 119640
tgttctgctg tgtgaactgg aacaatgtac cctctctgac ctcaatgtcc tctttccaaa 119700
ggggaactat tgctaccttt ctcagaaaag tagaaaggta cagagtcttg tataaaatcc 119760
aaactcaata aattctgatt tctgtcattc tttcttttca tgggtttggt cccgctcttc 119820
tgtaaaatgt gggacaattc tgatttagag atgtgggagt taggagttta taaaatgtgt 119880
tgcattgact ctccaacaaa acactctgga tgattccata cccctccctc ggcatttact 119940
gacaggctcc ctcagtagtg acccacagca cagccgggag tcctagcagc ctgaggggac 120000
tgctggttgg aacagggacg gaaaaggtct cccaaccacc atcactatca cctctcagca 120060
ccactgaggc ctcctggcct tgtcttttat tgagagactt tgttgtcata gcaacccaca 120120
gggtcatatc cccaaggccc cagagccaga gcaaaaagac agccaggaag agaggtttgc 120180
tgctgctgct gctgctgcta ccccactttt ctcatcacct gctttagatc tttctagctc 120240
cccctctgat gacctgactg tgcccctcaa gacaataaac ggaatgtagg ccacatcatc 120300
taccctgctc cttttacaaa ggaggggact gaggttcaga aataagagat gatttacccc 120360
agcttacaga ttttcttcat ggcaaagctg gaatgagaac ccaagtgttc tgactcctgt 120420
tctttcaaaa cccagcttct accggttatg ccaaaacatg acagaagttg ccgttggcaa 120480
ggcacaggca tgcctcagca taccctcccc tccagggctg ctgagtgggc aactctgccc 120540
acatttcctg gcaaggacaa tcaaggccca tcctgctttt tcccatgaga tgtttggagg 120600
agggcactgg ctctgcagta tattctcgtg atctggaatg acagccatcc ctcaggggac 120660
agataatgac cagaaccaca atggttattg cagcagtcag gtcagaaaat ttgagaggag 120720
ccctgctggc atccagtgaa gagtggccac accgaactga tttcacttct ctccttagac 120780
aacaaaatgc agcctgtgca ttctcctttc ttttttttt taattatact ttaagttctg 120840
gggtacatgt gcagaacata gagttttgtt acataggtat acacgtgcca tggcggtttg 120900
ctgcacccat caacccgtca tctacattag gtatttctcc taatgctatc cctcccctat 120960
ccctcacccc tgacaggctc cagtgtgtga tgttcctctc cctgtgtcca tgtgttctca 121020
ttgttcaact cccacttatg agtgagaaca tgcagtgttt ggttttctgt tcttgtgtta 121080
gtttgctgag aatgatggtt tgcatcctcc tttctttctg ctccactgtc ttgtccctct 121140

taatctcctt ctttcttctc ttccttattc cctggccctc tctctcccac tctaccttgg 121200
tgccctgcat tcaaattgac ctatgaggca gcccaaattg tttccccact attttctggc 121260
acgctggccc tggcccccac cagctgccca gaagacagct ggagtcccct tctagcggat 121320
gatgcctgtg gtgcgggttg ggcttgactt tctcatgaat gattatctga cttcttaccc 121380
gttctcttgc ctgtttatct tgccttcagc aggggatgag gatgaggatg aatccgggga 121440
ggagaggctg ccctcctgct ttgactacgt catgcacttc ctgactgtct tctggaaggt 121500
gctgtttgcc tgtgtgcccc ccacagagta ctgccacggc tgggcctgct tcgccgtctc 121560
catcctcatc attggcatgc tcaccgccat cattggggac ctggcctcgc acttcggctg 121620
caccattggt ctcaaagatt cagtcacagc tgttgttttc gtggcatttg gcacctctgt 121680
cccaggtgag agtgagaggt gcttgaattt gcaaagagga ttttacctgg ttcaaatgac 121740
ccctggactc catctcatta tcttccacac catctcagat ctgaacttaa cagagcctct 121800
gcccttaaag tgcacaaaag tcaatcaaag agatgaataa tgacattagt aatgacagct 121860
aatatttctt gagcactttc aatgtgacag acaccatgtg tgttcagcaa tttacacatt 121920
tacattttcc ccctgtaatg tttcccaaag ccctattaaa tagggtaagt tattatcccc 121980
acttcacaga caaagaaact gaggcccaca gaggttaagc tacatgccca agtaagtggt 122040
ccaatttctt aacctccaca ttatgtgagt agaccacaaa cagtgaaatt aaaagaatgt 122100
agatattgtt ctccttctat ttacctctgg cgatctctga gaggttaaag attagccagc 122160
tcaaagatat caaaggagaa atgcccacat acattcttgg cctcctctac ttggaaggac 122220
actgtgagta caaagtatct cctagcagga cagccaaagg aagttccaca gcttttatct 122280
ttttatagga tgaattacat actctttctt tttcttagga acactcagag acaaacagaa 122340
aggagcggac attcctttac tcattgaaca aatatttact gagcacctat tatgcctgtt 122400
acagtattgt gctagttttt gggactatag tgaaaggcaa gatacacatg cttccttctc 122460
cacgtggagt ttataatcta ctgaaggagg caactctcaa ctactgtaat taaagttatc 122520
ttgttaaatc ctaggaagaa aaagaaaagg tactgcatac ggaaggaagt tgggcctgaa 122580
tgtaggagtt agcaggtaga caggggctgc actagcccag gttctttact taattcagtt 122640
aggggctttg gggcctctga actctgaact tctgccaggg agctggcatc ccagttgccc 122700
cagaaagaaa cagagcacat cctcctgcag ggaagttagg ctgaatctca tcagacagga 122760
cttttctggc tgggccaagg gaaatctttc ctgtaccaag caaacatatc cttcaagaga 122820
gtagctgaat tcacatcaaa ttctaggaaa acctctttcc aaaaccccag cgcaggccag 122880
cggtattatt tgtccattag tgatgcaaga gatttagcta tcgtggaaat gcatcagaag 122940
gttggaaatt agatggatga tcccaggaag gcctgtggat gagatgccct gtgatctctg 123000
ttctccaagc cttgggggac ctgaactatc agaggggagg gaggaaatat gggggaaagc 123060
atagaggtgg gaagaaatat cagaggatca gaagcaaaaa acaacaataa caacagaaac 123120
aaaaacaaac aaacaaacaa aaaaacaagg ccataggcaa gaaagggtaa gaggttttct 123180
ctgggagatc taaaaaaaat ggcaataatg aggtaagcca ggcagatacc tttgggcatc 123240
tccaagtcct tgcaattggc caagacaaca gctaacaaca tttgaggctt taagaaggtt 123300
accctgtgat ccactcatct gatttagtgg ctttggctga agctctttgg atatagttga 123360
aggtacggaa agggtcctta catgaggact ttagggtcaa gtctcttgct aacatcctat 123420
gtgaccttgg gtaaattctt tgacccttat ttttcttacc tgtaaaataa aagaattggg 123480
ctagatgtct ctgacagtcc tccctgtatc tacaatctgt gccaagatct aaagtcaaac 123540
accctgcaag gccctgtgat acatatataa accacaaaga cagagccccg tcttccttga 123600
gtccacagtt caccctgcat gtccccatca tggttcccca acatgtcctc tgtccccaaa 123660
atccagcacc tcacccagtg ctcaatcagt aggcattgct caataactgt tggtggttcg 123720
tgaataaatg ccccatatga cagttaaaat caggcatcta ctccaagcag cttcccaggg 123780
tgtcaaggtt ccctggggag atattatggg atggcaaact tcccttactg aaaaagtagt 123840
caaaggagaa caataagccc actcagtaaa tatcagaact ggaaagccct tcagaatctt 123900
tcagatcact gcagatgagg aatgggaagc ccagactagg gatgtgacct acccagggcc 123960
acacggcttg cttgcggcag aactaggagt taggagtggc cccctagccc ttgtctctca 124020
ttcctgggtt cagcccacca gctcaagctg ctttttgggc atactggaag acaagccctg 124080
cacaccttag cctcctacca gttcccatgt gtctttgtcc ttttccagat acgtttgcca 124140
gcaaagctgc tgccctccag gatgtatatg cagacgcctc cattggcaac gtgacgggca 124200
gcaacgccgt caatgtcttc ctgggcatcg gcctggcctg gtccgtggcc gccatctact 124260
gggctctgca gggacaggag ttccacgtgt cggccggcac actggccttc tccgtcaccc 124320
tcttcaccat ctttgcattt gtctgcatca gcgtgctctt gtaccgaagg cggccgcacc 124380
tgggagggga gcttggtggc ccccgtggct gcaagctcgc cacaacatgg ctctttgtga 124440
gcctgtggct cctctacata ctctttgcca cactagaggc ctattgctac atcaaggggt 124500
tctaagccac acaacagagc ctccagcagg gcaggcctag gacttctcct aagagaaggg 124560


```
cacttcccca ccagtgatct ctcccgactg cactgccctg gagaggcagc atcaggacct 124620
aagccccagg aacttcaccc aacttaggcc ctggcaatta actgaaaggg caaagtctta 124680
atcaatcaaa caatggagga atcaccgact ttacacagta tttaattgaa tacaaacaag 124740
caacagcaac aaatccacct ccaccccatc tccccctcat atccctgacc caaagcaaag 124800
gtcagagcct ttcgcctcct tctattccat cttttgatta ttcctttgcc tctcatttct 124860
ttggaagcag ggtttctcct ctctgcccaa ttccatatgt ccctattatc tcactcagct 124920
gacaagacgt gaaaatgagt cacattcatg tggctggggt ggggttcttt tttcattgta 124980
atcattattg tggttgcttt cgttttgccg ttaggttttg cttattattt tgttttgtct 125040
tttttttctg aagtgagtga aaaaggtgcc acaaaggaat tccaggtccg agccaacaga 125100
gagaaacatg aatttttaga cacatgctct cctgccacct cttggctcca tcaagatcca 125160
gttccccatc tcactgtttt ctctgagttc ttgggaggag tgatggtgtt ggggtagaaa 125220
taagctcact cacccacgca gggtactaaa gatcttacag gagcttcaac tggagcagga 125280
ggagcttttt atgcttatgt tgaatcaagt cagatacaaa aagcaattgt ccctctttgc 125340
ccaagccttt ccaattctgt gtgtcttgtt gtgtcagtgt ccacttgtgt atccttctgc 125400
aggaagaccc gccaaataga agagatggga caaaaatagg aatggtgtgt gacgacaaag 125460
ggctactgga agaacaaaag ggatacaggc cttcttgatt atctttggct ttgtacctga 125520
ggcaggagag aagagatgtc caaccagtga gatctttaag agaaaagttt gtattttaaa 125580
tgtcaatgtg cctgagaaat gtcagcttca ccacgctctt gcttcctaat gctctataca 125640
aagagggctg actatatttc ttgaagtggt gtaaaaactt agagatttta taagagaacc 125700
aggggctccc ttcacctctc ctggtccctc aggtcacata tgaaagcatt tttacaagat 125760
aggaactgga attcctcatt tctcccatgt tcctgcttgt tcttaaactt catgaagcta 125820
tttttccagc ctatggggta gttcttgctc cagtaagagg aatcttagtt gtcataatcc 125880
cttggagcct gggtttttgg agaaagagat ctccgtgccc tacagacctt ttctcaacga 125940
atgtgggaag gacctggctt taaaacacgc acacaaacac acaaataaac agacataaga 126000
tgtcatcacg aaactgccca cggatcttta ggctttctgc attgacataa atacattttc 126060
taaggggggg ggggaagaaa ttaaaaaaca cctgttaatt ttaaacacat tttttaagaa 126120
aaaaataatt aaaaaagaaa cagtgctcat gtcataagct atgttgacag ttgccagtgg 126180
aaatgttggg ttggttcaaa aaaaaaataa aagctatact atatctctct acatacagct 126240
tgcttctacc tgtgtttctt cagtgaaagg tccaggggggc cactgtgggc ttcttgtgag 126300
gagacgtgac tcaggtgaag gtgtcacctc ctctcacact caggtgccaa tgtgtcagac 126360
ccagtatatt ctaagcaaaa atacttcagg aaaatgccac ttgtcaaaac ctggactttg 126420
cgaagttgga agatgtaagt agtagtaaaa gctgtggtaa ttatggagga aggaggtttc 126480
tgtatcagaa aggcattggc cgtgacagac tc                              126512
```

<210> 4
<211> 927
<212> PRT
<213> Rat

<400> 4

```
Met Ala Trp Leu Arg Leu Gln Pro Leu Thr Ser Ala Phe Leu His Phe
 1               5                  10                  15
Gly Leu Val Thr Phe Val Leu Phe Leu Asn Gly Leu Arg Ala Glu Ala
            20                  25                  30
Gly Asp Leu Arg Asp Val Pro Ser Ala Gly Gln Asn Asn Glu Ser Cys
        35                  40                  45
Ser Gly Ser Ser Asp Cys Lys Glu Gly Val Ile Leu Pro Ile Trp Tyr
    50                  55                  60
Pro Glu Asn Pro Ser Leu Gly Asp Lys Ile Ala Arg Val Ile Val Tyr
65                  70                  75                  80
Phe Val Ala Leu Ile Tyr Met Phe Leu Gly Val Ser Ile Ile Ala Asp
                85                  90                  95
Arg Phe Met Ala Ser Ile Glu Val Ile Thr Ser Gln Glu Arg Glu Val
            100                 105                 110
Thr Ile Lys Lys Pro Asn Gly Glu Thr Ser Thr Thr Thr Ile Arg Val
        115                 120                 125
Trp Asn Glu Thr Val Ser Asn Leu Thr Leu Met Ala Leu Gly Ser Ser
```

130 135 140
Ala Pro Glu Ile Leu Leu Ser Leu Ile Glu Val Cys Gly His Gly Phe
145 150 155 160
Ile Ala Gly Asp Leu Gly Pro Ser Thr Ile Val Gly Ser Ala Ala Phe
165 170 175
Asn Met Phe Ile Ile Ile Gly Ile Cys Val Tyr Val Ile Pro Asp Gly
180 185 190
Glu Thr Arg Lys Ile Lys His Leu Arg Val Phe Phe Val Thr Ala Ala
195 200 205
Trp Ser Val Phe Ala Tyr Ile Trp Leu Tyr Met Ile Leu Ala Val Phe
210 215 220
Ser Pro Gly Val Val Gln Val Trp Glu Gly Leu Leu Thr Leu Phe Phe
225 230 235 240
Phe Pro Val Cys Val Leu Leu Ala Trp Val Ala Asp Lys Arg Leu Leu
245 250 255
Phe Tyr Lys Tyr Met His Lys Arg Tyr Arg Thr Asp Lys His Arg Gly
260 265 270
Ile Ile Ile Glu Thr Glu Gly Glu His Pro Lys Gly Ile Glu Met Asp
275 280 285
Gly Lys Met Met Asn Ser His Phe Leu Asp Gly Asn Leu Ile Pro Leu
290 295 300
Glu Gly Lys Glu Val Asp Glu Ser Arg Arg Glu Met Ile Arg Ile Leu
305 310 315 320
Lys Asp Leu Lys Gln Lys His Pro Glu Lys Asp Leu Asp Gln Leu Val
325 330 335
Glu Met Ala Asn Tyr Tyr Ala Leu Ser His Gln Gln Lys Ser Arg Ala
340 345 350
Phe Tyr Arg Ile Gln Ala Thr Arg Met Met Thr Gly Ala Gly Asn Ile
355 360 365
Leu Lys Lys His Ala Ala Glu Gln Ala Lys Lys Thr Ala Ser Met Ser
370 375 380
Glu Val His Thr Asp Glu Pro Glu Asp Phe Ala Ser Lys Val Phe Phe
385 390 395 400
Asp Pro Cys Ser Tyr Gln Cys Leu Glu Asn Cys Gly Ala Val Leu Leu
405 410 415
Thr Val Val Arg Lys Gly Gly Asp Ile Ser Lys Thr Met Tyr Val Asp
420 425 430
Tyr Lys Thr Glu Asp Gly Ser Ala Asn Ala Gly Ala Asp Tyr Glu Phe
435 440 445
Thr Glu Gly Thr Val Val Leu Lys Pro Gly Glu Thr Gln Lys Glu Phe
450 455 460
Ser Val Gly Ile Ile Asp Asp Asp Ile Phe Glu Glu Asp Glu His Phe
465 470 475 480
Phe Val Arg Leu Ser Asn Val Arg Val Glu Glu Glu Gln Leu Glu Glu
485 490 495
Gly Met Thr Pro Ala Ile Leu Asn Ser Leu Pro Leu Pro Arg Ala Val
500 505 510
Leu Ala Ser Pro Cys Val Ala Thr Val Thr Ile Leu Asp Asp Asp His
515 520 525
Ala Gly Ile Phe Thr Phe Glu Cys Asp Thr Ile His Val Ser Glu Ser
530 535 540
Ile Gly Val Met Glu Val Lys Val Leu Arg Thr Ser Gly Ala Arg Gly
545 550 555 560
Thr Val Ile Val Pro Phe Arg Thr Val Glu Gly Thr Ala Lys Gly Gly
565 570 575
Gly Glu Asp Phe Glu Asp Thr Tyr Gly Glu Leu Glu Phe Lys Asn Asp
580 585 590

```
Glu Thr Val Lys Thr Ile Arg Val Lys Ile Val Asp Glu Glu Glu Tyr
        595                 600                 605
Glu Arg Gln Glu Asn Phe Phe Ile Ala Leu Gly Glu Pro Lys Trp Met
    610                 615                 620
Glu Arg Gly Ile Ser Ala Leu Leu Leu Ser Pro Glu Val Thr Asp Arg
625                 630                 635                 640
Lys Leu Thr Met Glu Glu Glu Glu Ala Lys Arg Ile Ala Glu Met Gly
                645                 650                 655
Lys Pro Val Leu Gly Glu His Pro Lys Leu Glu Val Ile Ile Glu Glu
            660                 665                 670
Ser Tyr Glu Phe Lys Ser Thr Val Asp Lys Leu Ile Lys Lys Thr Asn
        675                 680                 685
Leu Ala Leu Val Val Gly Thr His Ser Trp Arg Asp Gln Phe Met Glu
    690                 695                 700
Ala Ile Thr Val Ser Ala Ala Gly Asp Glu Glu Glu Asp Glu Ser Gly
705                 710                 715                 720
Glu Glu Arg Leu Pro Ser Cys Phe Asp Tyr Val Met His Phe Leu Thr
                725                 730                 735
Val Phe Trp Lys Val Leu Phe Ala Cys Val Pro Pro Thr Glu Tyr Cys
            740                 745                 750
His Gly Trp Ala Cys Phe Val Val Ser Ile Leu Ile Ile Gly Met Leu
        755                 760                 765
Thr Ala Ile Ile Gly Asp Leu Ala Ser His Phe Gly Cys Thr Ile Gly
    770                 775                 780
Leu Lys Asp Ser Val Thr Ala Val Val Phe Val Ala Phe Gly Thr Ser
785                 790                 795                 800
Val Pro Asp Thr Phe Ala Ser Lys Ala Ala Ala Leu Gln Asp Val Tyr
                805                 810                 815
Ala Asp Ala Ser Ile Gly Asn Val Thr Gly Ser Asn Ala Val Asn Val
            820                 825                 830
Phe Leu Gly Ile Gly Leu Ala Trp Ser Val Ala Ala Ile Tyr Trp Ala
        835                 840                 845
Met Gln Gly Gln Glu Phe His Val Ser Ala Gly Thr Leu Ala Phe Ser
    850                 855                 860
Val Thr Leu Phe Thr Ile Phe Ala Phe Val Cys Leu Ser Val Leu Leu
865                 870                 875                 880
Tyr Arg Arg Arg Pro His Leu Gly Gly Glu Leu Gly Gly Pro Arg Gly
                885                 890                 895
Cys Lys Leu Ala Thr Thr Trp Leu Phe Val Ser Leu Trp Leu Leu Tyr
            900                 905                 910
Val Leu Phe Ala Thr Leu Glu Ala Tyr Cys Tyr Ile Lys Gly Phe
        915                 920                 925
```